1 Vectors & Tensors

The mathematical modeling of the physical world requires knowledge of quite a few different mathematics subjects, such as Calculus, Differential Equations and Linear Algebra. These topics are usually encountered in fundamental mathematics courses. However, in a more thorough and in-depth treatment of mechanics, it is essential to describe the physical world using the concept of the **tensor**, and so we begin this book with a comprehensive chapter on the tensor.

The chapter is divided into three parts. The first part covers vectors (§1.1-1.7). The second part is concerned with second, and higher-order, tensors (§1.8-1.15). The second part covers much of the same ground as done in the first part, mainly generalizing the vector concepts and expressions to tensors. The final part (§1.16-1.19) (not required in the vast majority of applications) is concerned with generalizing the earlier work to curvilinear coordinate systems.

The first part comprises basic vector algebra, such as the dot product and the cross product; the mathematics of how the components of a vector transform between different coordinate systems; the symbolic, index and matrix notations for vectors; the differentiation of vectors, including the gradient, the divergence and the curl; the integration of vectors, including line, double, surface and volume integrals, and the integral theorems.

The second part comprises the definition of the tensor (and a re-definition of the vector); dyads and dyadics; the manipulation of tensors; properties of tensors, such as the trace, transpose, norm, determinant and principal values; special tensors, such as the spherical, identity and orthogonal tensors; the transformation of tensor components between different coordinate systems; the calculus of tensors, including the gradient of vectors and higher order tensors and the divergence of higher order tensors and special fourth order tensors.

In the first two parts, attention is restricted to rectangular Cartesian coordinates (except for brief forays into cylindrical and spherical coordinates). In the third part, curvilinear coordinates are introduced, including covariant and contravariant vectors and tensors, the metric coefficients, the physical components of vectors and tensors, the metric, coordinate transformation rules, tensor calculus, including the Christoffel symbols and covariant differentiation, and curvilinear coordinates for curved surfaces.

1.1 Vector Algebra

1.1.1 Scalars

A physical quantity which is completely described by a single real number is called a **scalar**. Physically, it is something which has a magnitude, and is completely described by this magnitude. Examples are **temperature, density** and **mass**. In the following, lowercase (usually Greek) letters, e.g. α, β, γ, will be used to represent scalars.

1.1.2 Vectors

The concept of the **vector** is used to describe physical quantities which have both a magnitude and a direction associated with them. Examples are **force**, **velocity**, **displacement** and **acceleration**.

Geometrically, a vector is represented by an arrow; the arrow defines the direction of the vector and the magnitude of the vector is represented by the length of the arrow, Fig. 1.1.1a.

Analytically, vectors will be represented by lowercase bold-face Latin letters, e.g. **a**, **r**, **q**.

The **magnitude** (or **length**) of a vector is denoted by $|\mathbf{a}|$ or a. It is a scalar and must be non-negative. Any vector whose length is 1 is called a **unit vector**; unit vectors will usually be denoted by **e**.

(a) (b)

Figure 1.1.1: (a) a vector; (b) addition of vectors

1.1.3 Vector Algebra

The operations of addition, subtraction and multiplication familiar in the algebra of numbers (or scalars) can be extended to an algebra of vectors.

The following definitions and properties fundamentally *define* the vector:

1. Sum of Vectors:

   The addition of vectors **a** and **b** is a vector **c** formed by placing the initial point of **b** on the terminal point of **a** and then joining the initial point of **a** to the terminal point of **b**. The sum is written $\mathbf{c} = \mathbf{a} + \mathbf{b}$. This definition is called the parallelogram law for vector addition because, in a geometrical interpretation of vector addition, **c** is the diagonal of a parallelogram formed by the two vectors **a** and **b**, Fig. 1.1.1b. The following properties hold for vector addition:

   $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ … commutative law

   $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$ … associative law

2. The Negative Vector:

   For each vector **a** there exists a **negative vector**. This vector has direction opposite to that of vector **a** but has the same magnitude; it is denoted by $-\mathbf{a}$. A geometrical interpretation of the negative vector is shown in Fig. 1.1.2a.

3. Subtraction of Vectors and the Zero Vector:

   The **subtraction** of two vectors **a** and **b** is defined by $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$, Fig. 1.1.2b. If $\mathbf{a} = \mathbf{b}$ then $\mathbf{a} - \mathbf{b}$ is defined as the **zero vector** (or **null vector**) and is represented by the symbol **o**. It has zero magnitude and unspecified direction. A **proper vector** is any vector other than the null vector. Thus the following properties hold:

   $$\mathbf{a} + \mathbf{o} = \mathbf{a}$$
   $$\mathbf{a} + (-\mathbf{a}) = \mathbf{o}$$

4. Scalar Multiplication:

   The product of a vector **a** by a scalar α is a vector $\alpha\mathbf{a}$ with magnitude $|\alpha|$ times the magnitude of **a** and with direction the same as or opposite to that of **a**, according as α is positive or negative. If $\alpha = 0$, $\alpha\mathbf{a}$ is the null vector. The following properties hold for scalar multiplication:

   $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$ … distributive law, over addition of scalars

   $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$ … distributive law, over addition of vectors

   $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$ … associative law for scalar multiplication

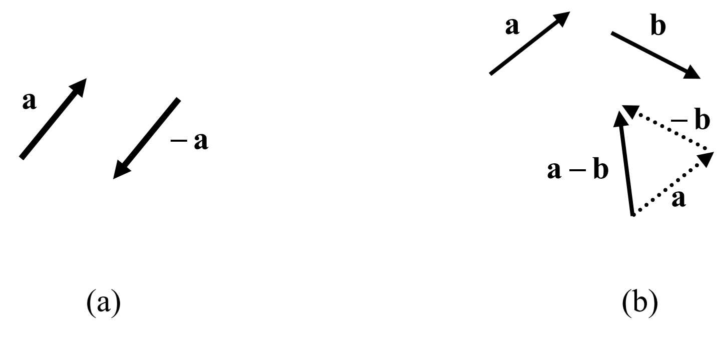


Figure 1.1.2: (a) negative of a vector; (b) subtraction of vectors

Note that when two vectors **a** and **b** are equal, they have the same direction and magnitude, regardless of the position of their initial points. Thus $\mathbf{a} = \mathbf{b}$ in Fig. 1.1.3. A particular position in space is not assigned here to a vector – it just has a magnitude and a direction. Such vectors are called **free**, to distinguish them from certain special vectors to which a particular position in space is actually assigned.

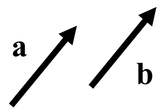


Figure 1.1.3: equal vectors

The vector as something with "magnitude and direction" and defined by the above rules is an element of one case of the mathematical structure, the **vector space**. The vector space is discussed in the next section, §1.2.

1.1.4 The Dot Product

The **dot product** of two vectors **a** and **b** (also called the **scalar product**) is denoted by $\mathbf{a} \cdot \mathbf{b}$. It is a scalar defined by

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta . \qquad (1.1.1)$$

θ here is the angle between the vectors when their initial points coincide and is restricted to the range $0 \le \theta \le \pi$, Fig. 1.1.4.

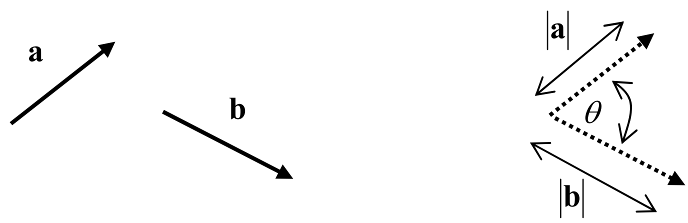


Figure 1.1.4: the dot product

An important property of the dot product is that if for two (proper) vectors **a** and **b**, the relation $\mathbf{a} \cdot \mathbf{b} = 0$, then **a** and **b** are perpendicular. The two vectors are said to be **orthogonal**. Also, $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}||\mathbf{a}|\cos(0)$, so that the length of a vector is $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$.

Another important property is that the **projection** of a vector **u** along the direction of a unit vector **e** is given by $\mathbf{u} \cdot \mathbf{e}$. This can be interpreted geometrically as in Fig. 1.1.5.

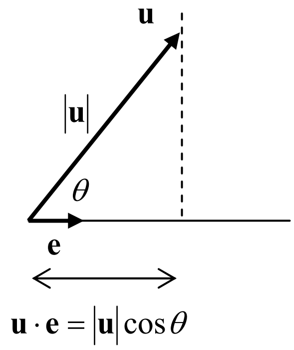


Figure 1.1.5: the projection of a vector along the direction of a unit vector

It follows that any vector **u** can be decomposed into a component parallel to a (unit) vector **e** and another component perpendicular to **e**, according to

$$\mathbf{u} = (\mathbf{u}\cdot\mathbf{e})\mathbf{e} + [\mathbf{u} - (\mathbf{u}\cdot\mathbf{e})\mathbf{e}] \tag{1.1.2}$$

The dot product possesses the following properties (which can be proved using the above definition) {▲Problem 6}:

(1) $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$ (commutative)
(2) $\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c}$ (distributive)
(3) $\alpha(\mathbf{a}\cdot\mathbf{b}) = \mathbf{a}\cdot(\alpha\mathbf{b})$
(4) $\mathbf{a}\cdot\mathbf{a} \geq 0$; and $\mathbf{a}\cdot\mathbf{a} = 0$ if and only if $\mathbf{a} = \mathbf{o}$

1.1.5 The Cross Product

The **cross product** of two vectors **a** and **b** (also called the **vector product**) is denoted by $\mathbf{a}\times\mathbf{b}$. It is a vector with magnitude

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta \tag{1.1.3}$$

with θ defined as for the dot product. It can be seen from the figure that the magnitude of $\mathbf{a}\times\mathbf{b}$ is equivalent to the area of the parallelogram determined by the two vectors **a** and **b**.

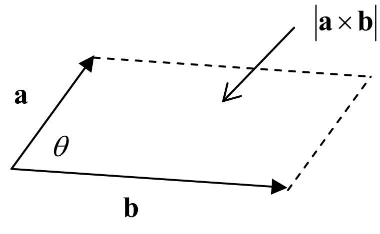


Figure 1.1.6: the magnitude of the cross product

The direction of this new vector is perpendicular to both **a** and **b**. Whether $\mathbf{a}\times\mathbf{b}$ points "up" or "down" is determined from the fact that the three vectors **a**, **b** and $\mathbf{a}\times\mathbf{b}$ form a **right handed system**. This means that if the thumb of the right hand is pointed in the direction of $\mathbf{a}\times\mathbf{b}$, and the open hand is directed in the direction of **a**, then the curling of the fingers of the right hand so that it closes should move the fingers through the angle θ, $0\le\theta\le\pi$, bringing them to **b**. Some examples are shown in Fig. 1.1.7.

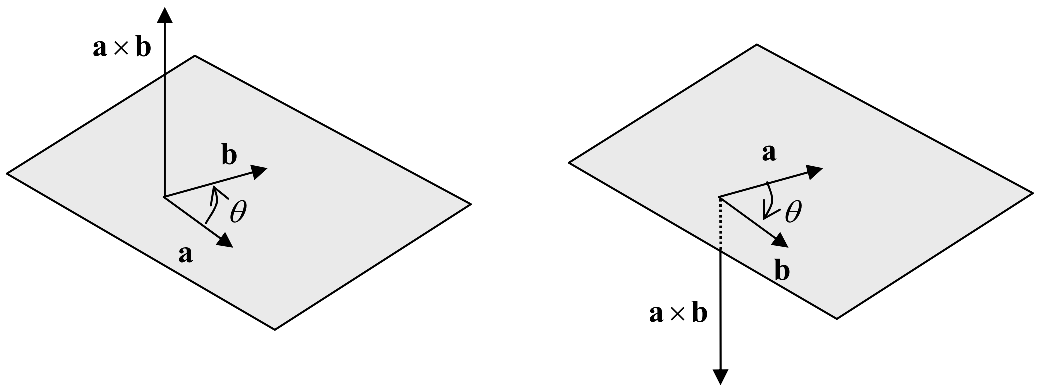


Figure 1.1.7: examples of the cross product

The cross product possesses the following properties (which can be proved using the above definition):

(1) $\mathbf{a}\times\mathbf{b}=-\mathbf{b}\times\mathbf{a}$ (*not* commutative)
(2) $\mathbf{a}\times(\mathbf{b}+\mathbf{c})=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}$ (distributive)
(3) $\alpha(\mathbf{a}\times\mathbf{b})=\mathbf{a}\times(\alpha\mathbf{b})$
(4) $\mathbf{a}\times\mathbf{b}=\mathbf{o}$ if and only if **a** and **b** $(\neq\mathbf{o})$ are parallel ("linearly dependent")

The Triple Scalar Product

The **triple scalar product**, or **box product**, of three vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ is defined by

$$\boxed{(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}=(\mathbf{v}\times\mathbf{w})\cdot\mathbf{u}=(\mathbf{w}\times\mathbf{u})\cdot\mathbf{v}} \quad \textbf{Triple Scalar Product} \qquad (1.1.4)$$

Its importance lies in the fact that, if the three vectors form a right-handed triad, then the volume V of a parallelepiped spanned by the three vectors is equal to the box product.

To see this, let **e** be a unit vector in the direction of $\mathbf{u}\times\mathbf{v}$, Fig. 1.1.8. Then the projection of **w** on $\mathbf{u}\times\mathbf{v}$ is $h=\mathbf{w}\cdot\mathbf{e}$, and

$$\begin{aligned}\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) &= \mathbf{w}\cdot(|\mathbf{u}\times\mathbf{v}|\mathbf{e})\\ &= |\mathbf{u}\times\mathbf{v}|h\\ &= V\end{aligned} \qquad (1.1.5)$$

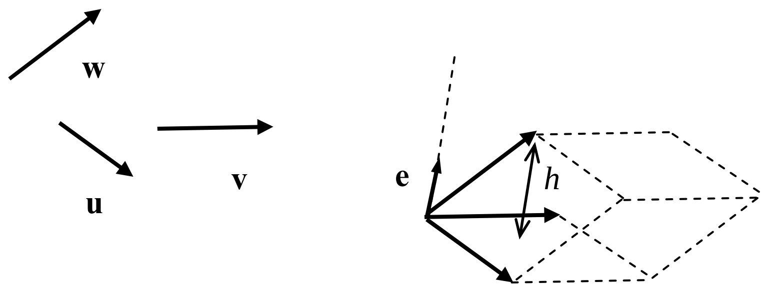


Figure 1.1.8: the triple scalar product

Note: if the three vectors do not form a right handed triad, then the triple scalar product yields the negative of the volume. For example, using the vectors above, $(\mathbf{w}\times\mathbf{v})\cdot\mathbf{u} = -V$.

1.1.6 Vectors and Points

Vectors are objects which have magnitude and direction, but they do not have any specific location in space. On the other hand, a **point** has a certain position in space, and the only characteristic that distinguishes one point from another is its position. Points cannot be "added" together like vectors. On the other hand, a vector **v** can be added to a point **p** to give a new point **q**, $\mathbf{q} = \mathbf{v} + \mathbf{p}$, Fig. 1.1.9. Similarly, the "difference" between two points gives a vector, $\mathbf{q} - \mathbf{p} = \mathbf{v}$. Note that the notion of point as defined here is slightly different to the familiar point in space with axes and origin – the concept of origin is not necessary for these points and their simple operations with vectors.

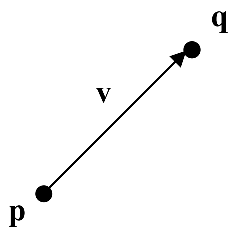
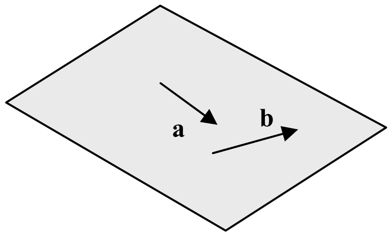


Figure 1.1.9: adding vectors to points

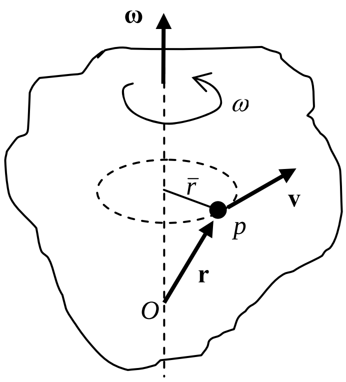
1.1.7 Problems

1. Which of the following are scalars and which are vectors?
   (i) weight
   (ii) specific heat
   (iii) momentum
   (iv) energy
   (v) volume
2. Find the magnitude of the sum of three unit vectors drawn from a common vertex of a cube along three of its sides.

3. Consider two **non-collinear** (not parallel) vectors **a** and **b**. Show that a vector **r** lying in the same plane as these vectors can be written in the form $\mathbf{r} = p\mathbf{a} + q\mathbf{b}$, where *p* and *q* are scalars. [Note: one says that all the vectors **r** in the plane are specified by the **base** vectors **a** and **b**.]
4. Show that the dot product of two vectors **u** and **v** can be interpreted as the magnitude of **u** times the component of **v** in the direction of **u**.
5. The work done by a force, represented by a vector **F**, in moving an object a given distance is the product of the component of force in the given direction times the distance moved. If the vector **s** represents the direction and magnitude (distance) the object is moved, show that the work done is equivalent to $\mathbf{F} \cdot \mathbf{s}$.
6. Prove that the dot product is commutative, $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$. [Note: this is equivalent to saying, for example, that the work done in problem 5 is also equal to the component of **s** in the direction of the force, times the magnitude of the force.]
7. Sketch $\mathbf{b} \times \mathbf{a}$ if **a** and **b** are as shown below.



8. Show that $|\mathbf{a} \times \mathbf{b}|^2 + |\mathbf{a} \cdot \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2$.
9. Suppose that a rigid body rotates about an axis *O* with angular speed ω, as shown below. Consider a point *p* in the body with position vector **r**. Show that the velocity **v** of *p* is given by $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, where $\boldsymbol{\omega}$ is the vector with magnitude ω and whose direction is that in which a right-handed screw would advance under the rotation. [Note: let *s* be the arc-length traced out by the particle as it rotates through an angle θ on a circle of radius $\bar{r}$, then $v = |\mathbf{v}| = \bar{r}\omega$ (since $s = \bar{r}\theta, ds/dt = \bar{r}(d\theta/dt)$).]



10. Show, geometrically, that the dot and cross in the triple scalar product can be interchanged: $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.
11. Show that the **triple vector product** $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ lies in the plane spanned by the vectors **a** and **b**.

1.2 Vector Spaces

The notion of the vector presented in the previous section is here re-cast in a more formal and abstract way, using some basic concepts of Linear Algebra and Topology. This might seem at first to be unnecessarily complicating matters, but this approach turns out to be helpful in unifying and bringing clarity to much of the theory which follows.

Some background theory which complements this material is given in Appendix A to this Chapter, §1.A.

1.2.1 The Vector Space

The vectors introduced in the previous section obey certain rules, those listed in §1.1.3. It turns out that many other mathematical objects obey the same list of rules. For that reason, the mathematical structure defined by these rules is given a special name, the **linear space** or **vector space**.

First, a **set** is any well-defined list, collection, or class of objects, which could be finite or infinite. An example of a set might be

$$B = \{x \mid x \le 3\} \tag{1.2.1}$$

which reads "B is the set of objects x such that x satisfies the property $x \le 3$". Members of a set are referred to as **elements**.

Consider now the **field**[1] of real numbers R. The elements of R are referred to as **scalars**. Let V be a non-empty set of elements $\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$ with rules of **addition** and **scalar multiplication**, that is there is a **sum** $\mathbf{a} + \mathbf{b} \in V$ for any $\mathbf{a}, \mathbf{b} \in V$ and a **product** $\alpha\mathbf{a} \in V$ for any $\mathbf{a} \in V$, $\alpha \in R$. Then V is called a **(real)**[2] **vector space** over R if the following eight axioms hold:

1. *associative law for addition*: for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$, one has $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
2. *zero element*: there exists an element $\mathbf{o} \in V$, called the zero element, such that $\mathbf{a} + \mathbf{o} = \mathbf{o} + \mathbf{a} = \mathbf{a}$ for every $\mathbf{a} \in V$
3. *negative* (or *inverse*): for each $\mathbf{a} \in V$ there exists an element $-\mathbf{a} \in V$, called the negative of $\mathbf{a}$, such that $\mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{o}$
4. *commutative law for addition*: for any $\mathbf{a}, \mathbf{b} \in V$, one has $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
5. *distributive law, over addition of elements of V*: for any $\mathbf{a}, \mathbf{b} \in V$ and scalar $\alpha \in R$, $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$
6. *distributive law, over addition of scalars*: for any $\mathbf{a} \in V$ and scalars $\alpha, \beta \in R$, $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$

[1] A **field** is another mathematical structure (see Appendix A to this Chapter, §1.A). For example, the set of complex numbers is a field. In what follows, the only field which will be used is the familiar set of real numbers with the usual operations of addition and multiplication.

[2] "real", since the associated field is the reals. The word *real* will usually be omitted in what follows for brevity.

7. *associative law for multiplication*: for any $\mathbf{a} \in V$ and scalars $\alpha, \beta \in R$, $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$
8. *unit multiplication*: for the unit scalar $1 \in R$, $1\mathbf{a} = \mathbf{a}$ for any $\mathbf{a} \in V$.

The set of vectors as objects with “magnitude and direction” discussed in the previous section satisfy these rules and therefore form a vector space over *R*. However, despite the name “vector” space, other objects, which are *not* the familiar geometric vectors, can also form a vector space over *R*, as will be seen in a later section.

1.2.2 Inner Product Space

Just as the vector of the previous section is an element of a vector space, next is introduced the notion that the vector dot product is one example of the more general **inner product**.

First, a **function** (or **mapping**) is an assignment which assigns to *each* element of a set *A* a *unique* element of a set B, and is denoted by

$$f : A \to B \tag{1.2.2}$$

An **ordered pair** (a, b) consists of two elements *a* and *b* in which one of them is designated the first element and the other is designated the second element The **product set** (or **Cartesian product**) $A \times B$ consists of all ordered pairs (a, b) where $a \in A$ and $b \in B$:

$$A \times B = \{(a, b) \mid a \in A, b \in B\} \tag{1.2.3}$$

Now let *V* be a real vector space. An **inner product** (or **scalar product**) on *V* is a mapping that associates to each ordered pair of elements **x**, **y**, a scalar, denoted by $\langle x, y \rangle$,

$$\langle \cdot, \cdot \rangle : V \times V \to R \tag{1.2.4}$$

that satisfies the following properties, for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\alpha \in R$:

1. *additivity*: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
2. *homogeneity*: $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$
3. *symmetry*: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
4. *positive definiteness*: $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ when $\mathbf{x} \neq \mathbf{o}$

From these properties, it follows that, if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in V$, then $\mathbf{x} = 0$

A vector space with an associated inner product is called an **inner product space**.

Two elements of an inner product space are said to be **orthogonal** if

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \tag{1.2.5}$$

and a set of elements of *V*, $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots\}$, are said to form an **orthogonal set** if every element in the set is orthogonal to every other element:

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0, \quad \langle \mathbf{x}, \mathbf{z} \rangle = 0, \quad \langle \mathbf{y}, \mathbf{z} \rangle = 0, \quad \text{etc.} \tag{1.2.6}$$

The above properties are those listed in §1.1.4, and so the set of vectors with the associated dot product forms an inner product space.

Euclidean Vector Space

The set of real triplets (x_1, x_2, x_3) under the usual rules of addition and multiplication forms a vector space R^3. With the inner product defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

one has the inner product space known as (three dimensional) **Euclidean vector space**, and denoted by *E*. This inner product allows one to take distances (and angles) between elements of *E* through the norm (length) and metric (distance) concepts discussed next.

1.2.3 Normed Space

Let *V* be a real vector space. A **norm** on *V* is a real-valued function,

$$\| \ \| : V \to R \tag{1.2.7}$$

that satisfies the following properties, for $\mathbf{x}, \mathbf{y} \in V$, $\alpha \in R$:

1. *positivity*: $\|\mathbf{x}\| \geq 0$
2. *triangle inequality*: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
3. *homogeneity*: $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
4. *positive definiteness*: $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{o}$

A vector space with an associated norm is called a **normed vector space**. Many different norms can be defined on a given vector space, each one giving a different normed linear space. A natural norm for the inner product space is

$$\|\mathbf{x}\| \equiv \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \tag{1.2.8}$$

It can be seen that this norm indeed satisfies the defining properties. When the inner product is the vector dot product, the norm defined by 1.2.8 is the familiar vector "length".

One important consequence of the definitions of inner product and norm is the **Schwarz inequality**, which states that

$$\left|\langle \mathbf{x}, \mathbf{y} \rangle\right| \le \|\mathbf{x}\|\|\mathbf{y}\| \tag{1.2.9}$$

One can now define the **angle** between two elements of *V* to be

$$\theta : V \times V \to R, \qquad \theta(\mathbf{x}, \mathbf{y}) \equiv \cos^{-1}\left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}\right) \tag{1.2.10}$$

The quantity inside the curved brackets here is necessarily between -1 and $+1$, by the Schwarz inequality, and hence the angle θ is indeed a real number.

1.2.4 Metric Spaces

Metric spaces are built on the concept of "distance" between objects. This is a generalization of the familiar distance between two points on the real line.

Consider a set *X*. A **metric** is a real valued function,

$$d(\cdot, \cdot) : X \times X \to R \tag{1.2.11}$$

that satisfies the following properties, for $\mathbf{x}, \mathbf{y} \in X$:

1. positive: $d(\mathbf{x}, \mathbf{y}) \ge 0$ and $d(\mathbf{x}, \mathbf{x}) = 0$, for all $\mathbf{x}, \mathbf{y} \in X$
2. strictly positive: if $d(\mathbf{x}, \mathbf{y}) = 0$ then $\mathbf{x} = \mathbf{y}$, for all $\mathbf{x}, \mathbf{y} \in X$
3. symmetry: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, for all $\mathbf{x}, \mathbf{y} \in X$
4. triangle inequality: $d(\mathbf{x}, \mathbf{y}) \le d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$, for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$

A set *X* with an associated metric is called a **metric space**. The set *X* can have more than one metric defined on it, with different metrics producing different metric spaces.

Consider now a normed vector space. This space naturally has a metric defined on it:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \tag{1.2.12}$$

and thus the normed vector space *is* a metric space. For the set of vectors with the dot product, this gives the "distance" between two vectors $\mathbf{x}, \mathbf{y}$.

1.2.5 The Affine Space

Consider a set *P*, the elements of which are called **points**. Consider also an associated vector space *V*. *P* is an **affine space** when:

(i) given two points $p, q \in P$, one can define a **difference**, $q - p$ which is a unique element **v** of *V*, i.e. $\mathbf{v} \equiv \mathbf{v}(q, p) = q - p \in V$ (called a **translation vector**),

(ii) given a point $p \in P$ and $\mathbf{v} \in V$, one can define the **sum** $\mathbf{v} + p$ which is a unique point *q* of *P*, i.e. $q = \mathbf{v} + p \in P$,

and for which the following property holds: for $p, q, r \in P$:

$$(q - r) + (r - p) = (q - p)$$

From the above, one has for the affine space that $p - p = \mathbf{o}$ and $q - p = -(p - q)$, for all $p, q \in P$.

One can take the sum of vectors, according to the structure of the vector space, but one cannot take the sum of points, only the difference between two points.

A key point is that there is no notion of **origin** in the affine space. There is no special or significant point in the set *P*, unlike with the vector space, where there is a special zero element, **o**, which has its own axiom (see axiom 2 in §1.2.1 above).

Suppose now that the associated vector space is a Euclidean vector space, i.e. an inner product space. Define the **distance** between two points through the inner product associated with *V*,

$$d(p, q) = \|q - p\| = \sqrt{\langle q - p, q - p \rangle} \tag{1.2.13}$$

It can be shown that this mapping $d : P \times P \to R$ is a metric, i.e. it satisfies the metric properties, and thus *P* is a metric space (although it is not a vector space). In this case, *P* is referred to as **Euclidean point space**, **Euclidean affine space** or, simply, **Euclidean space**.

Whereas in Euclidean vector space there is a zero element, in Euclidean point space there is none – apart from that, the two spaces are the same and, apart from certain special cases, one does not need to distinguish between them.

Note: one can generalise the simple affine space into a vector space by choosing some fixed $o \in P$ to be an origin. In that case, $\mathbf{v} \equiv \mathbf{v}(p, o) = p - o$ is called the **position vector** of *p* relative to *o*. Then one can define the sum of two points through $p + q = o + (\mathbf{v} + \mathbf{w})$, where $\mathbf{v} = p - o$, $\mathbf{w} = q - o$.[3]

[3] One also has to define a scaling, e.g. $\alpha p \equiv 0 + \alpha \mathbf{v}$, where α is in the associated field (of real numbers).

1.3 Cartesian Vectors

So far the discussion has been in **symbolic notation**[1], that is, no reference to 'axes' or 'components' or 'coordinates' is made, implied or required. The vectors exist independently of any coordinate system. It turns out that much of vector (tensor) mathematics is more concise and easier to manipulate in such notation than in terms of corresponding component notations. However, there are many circumstances in which use of the component forms of vectors (and tensors) is more helpful – or essential. In this section, vectors are discussed in terms of components – **component form**.

1.3.1 The Cartesian Basis

Consider three dimensional (Euclidean) space. In this space, consider the three unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ having the properties

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_2 \cdot \mathbf{e}_3 = \mathbf{e}_3 \cdot \mathbf{e}_1 = 0, \qquad (1.3.1)$$

so that they are mutually perpendicular (mutually **orthogonal**), and

$$\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_3 \cdot \mathbf{e}_3 = 1, \qquad (1.3.2)$$

so that they are unit vectors. Such a set of orthogonal unit vectors is called an **orthonormal** set, Fig. 1.3.1. Note further that this orthonormal system $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is **right-handed**, by which is meant $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$ (or $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ or $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$).

This set of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ forms a **basis,** by which is meant that any other vector can be written as a **linear combination** of these vectors, i.e. in the form

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 \qquad (1.3.3)$$

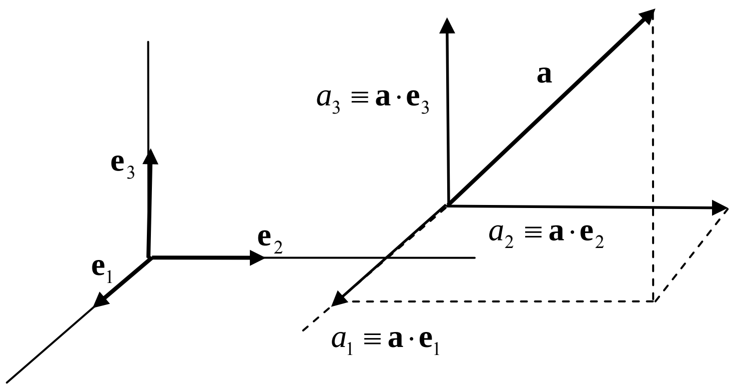


Figure 1.3.1: an orthonormal set of base vectors and Cartesian components

[1] or **absolute** or **invariant** or **direct** or **vector** notation

By repeated application of Eqn. 1.1.2 to a vector **a**, and using 1.3.2, the scalars in 1.3.3 can be expressed as (see Fig. 1.3.1)

$$a_1 = \mathbf{a}\cdot\mathbf{e}_1, \quad a_2 = \mathbf{a}\cdot\mathbf{e}_2, \quad a_3 = \mathbf{a}\cdot\mathbf{e}_3 \tag{1.3.4}$$

The scalars a_1, a_2 and a_3 are called the **Cartesian components** of **a** in the given basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The unit vectors are called **base vectors** when used for this purpose.

Note that it is not necessary to have three mutually orthogonal vectors, or vectors of unit size, or a right-handed system, to form a basis – only that the three vectors are not co-planar. The right-handed orthonormal set is often the easiest basis to use in practice, but this is not always the case – for example, when one wants to describe a body with curved boundaries (e.g., see §1.6.10).

The dot product of two vectors **u** and **v**, referred to the above basis, can be written as

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= (u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3)\cdot(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3)\\ &= u_1v_1(\mathbf{e}_1\cdot\mathbf{e}_1) + u_1v_2(\mathbf{e}_1\cdot\mathbf{e}_2) + u_1v_3(\mathbf{e}_1\cdot\mathbf{e}_3)\\ &\quad + u_2v_1(\mathbf{e}_2\cdot\mathbf{e}_1) + u_2v_2(\mathbf{e}_2\cdot\mathbf{e}_2) + u_2v_3(\mathbf{e}_2\cdot\mathbf{e}_3)\\ &\quad\quad + u_3v_1(\mathbf{e}_3\cdot\mathbf{e}_1) + u_3v_2(\mathbf{e}_3\cdot\mathbf{e}_2) + u_3v_3(\mathbf{e}_3\cdot\mathbf{e}_3)\\ &= u_1v_1 + u_2v_2 + u_3v_3\end{aligned} \tag{1.3.5}$$

Similarly, the cross product is

$$\begin{aligned}\mathbf{u}\times\mathbf{v} &= (u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3)\times(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3)\\ &= u_1v_1(\mathbf{e}_1\times\mathbf{e}_1) + u_1v_2(\mathbf{e}_1\times\mathbf{e}_2) + u_1v_3(\mathbf{e}_1\times\mathbf{e}_3)\\ &\quad + u_2v_1(\mathbf{e}_2\times\mathbf{e}_1) + u_2v_2(\mathbf{e}_2\times\mathbf{e}_2) + u_2v_3(\mathbf{e}_2\times\mathbf{e}_3)\\ &\quad\quad + u_3v_1(\mathbf{e}_3\times\mathbf{e}_1) + u_3v_2(\mathbf{e}_3\times\mathbf{e}_2) + u_3v_3(\mathbf{e}_3\times\mathbf{e}_3)\\ &= (u_2v_3 - u_3v_2)\mathbf{e}_1 - (u_1v_3 - u_3v_1)\mathbf{e}_2 + (u_1v_2 - u_2v_1)\mathbf{e}_3\end{aligned} \tag{1.3.6}$$

This is often written in the form

$$\mathbf{u}\times\mathbf{v} = \begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix}, \tag{1.3.7}$$

that is, the cross product is equal to the determinant of the 3×3 matrix

$$\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{bmatrix}$$

1.3.2 The Index Notation

The expression for the cross product in terms of components, Eqn. 1.3.6, is quite lengthy – for more complicated quantities things get unmanageably long. Thus a short-hand notation is used for these component equations, and this **index notation**[2] is described here.

In the index notation, the expression for the vector **a** in terms of the components a_1, a_2, a_3 and the corresponding basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is written as

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 = \sum_{i=1}^{3} a_i\mathbf{e}_i \qquad (1.3.8)$$

This can be simplified further by using Einstein's **summation convention**, whereby the summation sign is dropped and it is understood that for a repeated index (*i* in this case) a summation over the range of the index (3 in this case[3]) is implied. Thus one writes $\mathbf{a} = a_i\mathbf{e}_i$. This can be further shortened to, simply, a_i.

The dot product of two vectors written in the index notation reads

$$\boxed{\mathbf{u}\cdot\mathbf{v} = u_i v_i} \qquad \textbf{Dot Product} \qquad (1.3.9)$$

The repeated index *i* is called a **dummy index**, because it can be replaced with any other letter and the sum is the same; for example, this could equally well be written as $\mathbf{u}\cdot\mathbf{v} = u_j v_j$ or $u_k v_k$.

For the purpose of writing the vector cross product in index notation, the **permutation symbol** (or **alternating symbol**) ε_{ijk} can be introduced:

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0 & \text{if two or more indices are equal} \end{cases} \qquad (1.3.10)$$

For example (see Fig. 1.3.2),

$$\begin{aligned} \varepsilon_{123} &= +1 \\ \varepsilon_{132} &= -1 \\ \varepsilon_{122} &= 0 \end{aligned}$$

[2] or **indicial** or **subscript** or **suffix** notation
[3] 2 in the case of a two-dimensional space/analysis

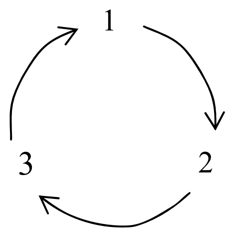


Figure 1.3.2: schematic for the permutation symbol (clockwise gives +1)

Note that

$$\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij} = -\varepsilon_{jik} = -\varepsilon_{kji} = -\varepsilon_{ikj} \tag{1.3.11}$$

and that, in terms of the base vectors {▲Problem 7},

$$\mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ijk}\mathbf{e}_k \tag{1.3.12}$$

and {▲Problem 7}

$$\varepsilon_{ijk} = \left(\mathbf{e}_i \times \mathbf{e}_j\right)\cdot \mathbf{e}_k\,. \tag{1.3.13}$$

The cross product can now be written concisely as {▲Problem 8}

$$\boxed{\mathbf{u}\times\mathbf{v} = \varepsilon_{ijk} u_i v_j \mathbf{e}_k} \qquad \textbf{Cross Product} \tag{1.3.14}$$

Introduce next the **Kronecker delta symbol** δ_{ij}, defined by

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{1.3.15}$$

Note that $\delta_{11} = 1$ but, using the index notation, $\delta_{ii} = 3$. The Kronecker delta allows one to write the expressions defining the orthonormal basis vectors (1.3.1, 1.3.2) in the compact form

$$\boxed{\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}} \qquad \textbf{Orthonormal Basis Rule} \tag{1.3.16}$$

The triple scalar product (1.1.4) can now be written as

$$\begin{aligned}(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w} &= \left(\varepsilon_{ijk} u_i v_j \mathbf{e}_k\right)\cdot w_m \mathbf{e}_m \\ &= \varepsilon_{ijk} u_i v_j w_m \delta_{km} \\ &= \varepsilon_{ijk} u_i v_j w_k \\ &= \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}\end{aligned} \tag{1.3.17}$$

Note that, since the determinant of a matrix is equal to the determinant of the transpose of a matrix, this is equivalent to

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} \tag{1.3.18}$$

Here follow some useful formulae involving the permutation and Kronecker delta symbol {▲Problem 13}:

$$\begin{aligned} \varepsilon_{ijk}\varepsilon_{kpq} &= \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp} \\ \varepsilon_{ijk}\varepsilon_{ijp} &= 2\delta_{pk} \end{aligned} \tag{1.3.19}$$

Finally, here are some other important identities involving vectors; the third of these is called **Lagrange's identity**[4] {▲Problem 15}:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) &= |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 \\ \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \\ (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{a} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix} \\ (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{d})]\mathbf{c} - [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\mathbf{d} \\ [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\mathbf{d} &= [\mathbf{d} \cdot (\mathbf{b} \times \mathbf{c})]\mathbf{a} + [\mathbf{a} \cdot (\mathbf{d} \times \mathbf{c})]\mathbf{b} + [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{d})]\mathbf{c} \end{aligned} \tag{1.3.20}$$

1.3.3 Matrix Notation for Vectors

The symbolic notation $\mathbf{v}$ and index notation $v_i\mathbf{e}_i$ (or simply v_i) can be used to denote a vector. Another notation is the **matrix notation**: the vector **v** can be represented by a 3×1 matrix (a **column vector**):

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Matrices will be denoted by square brackets, so a shorthand notation for this matrix/vector would be $[\mathbf{v}]$. The elements of the matrix $[\mathbf{v}]$ can be written in the **element form** v_i. The element form for a matrix is essentially the same as the index notation for the vector it represents.

[4] to be precise, the special case of 1.3.20c, 1.3.20a, is Lagrange's identity

Formally, a vector can be represented by the ordered triplet of real numbers, (v_1, v_2, v_3). The set of all vectors can be represented by R^3, the set of all ordered triplets of real numbers:

$$R^3 = \{(v_1, v_2, v_3) \mid v_1, v_2, v_3 \in R\} \tag{1.3.21}$$

It is important to *note the distinction between a vector and a matrix*: the former is a mathematical object independent of any basis, the latter is a representation of the vector with respect to a particular basis – use a different set of basis vectors and the elements of the matrix will change, but the matrix is still describing the same vector. Said another way, there is a difference between an element (vector) **v** of Euclidean vector space and an ordered triplet $v_i \in R^3$. This notion will be discussed more fully in the next section.

As an example, the dot product can be written in the matrix notation as

$$[\mathbf{u}^{\mathrm{T}}][\mathbf{v}] = [u_1 \quad u_2 \quad u_3]\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

“short” matrix notation

“full” matrix notation

Here, the notation $[\mathbf{u}^{\mathrm{T}}]$ denotes the 1×3 matrix (the **row vector**). The result is a 1×1 matrix, i.e. a scalar, in element form $u_i v_i$.

1.3.4 Cartesian Coordinates

Thus far, the notion of an origin has not been used. Choose a point **o** in Euclidean (point) space, to be called the **origin**. An origin together with a right-handed orthonormal basis $\{\mathbf{e}_i\}$ constitutes a (**rectangular**) **Cartesian coordinate system**, Fig. 1.3.3.

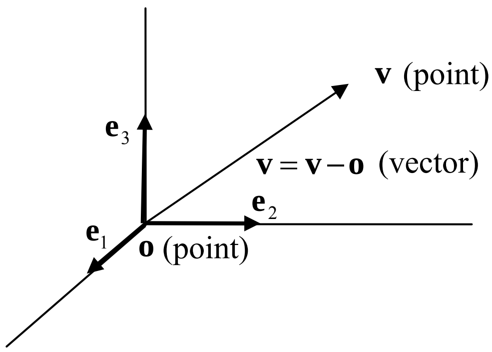


Figure 1.3.3: a Cartesian coordinate system

A second point **v** then defines a **position vector** $\mathbf{v}-\mathbf{o}$, Fig. 1.3.3. The components of the vector $\mathbf{v}-\mathbf{o}$ are called the (**rectangular**) **Cartesian coordinates** of the point **v** [5]. For brevity, the vector $\mathbf{v}-\mathbf{o}$ is simply labelled **v**, that is, one uses the same symbol for both the position vector and associated point.

1.3.5 Problems

1. Evaluate $\mathbf{u}\cdot\mathbf{v}$ where $\mathbf{u}=\mathbf{e}_1+3\mathbf{e}_2-2\mathbf{e}_3$, $\mathbf{v}=4\mathbf{e}_1-2\mathbf{e}_2+4\mathbf{e}_3$.
2. Prove that for any vector **u**, $\mathbf{u}=(\mathbf{u}\cdot\mathbf{e}_1)\mathbf{e}_1+(\mathbf{u}\cdot\mathbf{e}_2)\mathbf{e}_2+(\mathbf{u}\cdot\mathbf{e}_3)\mathbf{e}_3$. [Hint: write **u** in component form.]
3. Find the projection of the vector $\mathbf{u}=\mathbf{e}_1-2\mathbf{e}_2+\mathbf{e}_3$ on the vector $\mathbf{v}=4\mathbf{e}_1-4\mathbf{e}_2+7\mathbf{e}_3$.
4. Find the angle between $\mathbf{u}=3\mathbf{e}_1+2\mathbf{e}_2-6\mathbf{e}_3$ and $\mathbf{v}=4\mathbf{e}_1-3\mathbf{e}_2+\mathbf{e}_3$.
5. Write down an expression for a unit vector parallel to the resultant of two vectors **u** and **v** (in symbolic notation). Find this vector when $\mathbf{u}=2\mathbf{e}_1+4\mathbf{e}_2-5\mathbf{e}_3$, $\mathbf{v}=\mathbf{e}_1+2\mathbf{e}_2+3\mathbf{e}_3$ (in component form). Check that your final vector is indeed a unit vector.
6. Evaluate $\mathbf{u}\times\mathbf{v}$, where $\mathbf{u}=-\mathbf{e}_1-2\mathbf{e}_2+2\mathbf{e}_3$, $\mathbf{v}=2\mathbf{e}_1-2\mathbf{e}_2+\mathbf{e}_3$.
7. Verify that $\mathbf{e}_i\times\mathbf{e}_j=\varepsilon_{ijm}\mathbf{e}_m$. Hence, by dotting each side with $\mathbf{e}_k$, show that $\varepsilon_{ijk}=(\mathbf{e}_i\times\mathbf{e}_j)\cdot\mathbf{e}_k$.
8. Show that $\mathbf{u}\times\mathbf{v}=\varepsilon_{ijk}u_iv_j\mathbf{e}_k$.
9. The triple scalar product is given by $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}=\varepsilon_{ijk}u_iv_jw_k$. Expand this equation and simplify, so as to express the triple scalar product in full (non-index) component form.
10. Write the following in index notation: $|\mathbf{v}|$, $\mathbf{v}\cdot\mathbf{e}_1$, $\mathbf{v}\cdot\mathbf{e}_k$.
11. Show that $\delta_{ij}a_ib_j$ is equivalent to $\mathbf{a}\cdot\mathbf{b}$.
12. Verify that $\varepsilon_{ijk}\varepsilon_{ijk}=6$.
13. Verify that $\varepsilon_{ijk}\varepsilon_{kpq}=\delta_{ip}\delta_{jq}-\delta_{iq}\delta_{jp}$ and hence show that $\varepsilon_{ijk}\varepsilon_{ijp}=2\delta_{pk}$.
14. Evaluate or simplify the following expressions:
    (a) δ_{kk} (b) $\delta_{ij}\delta_{ij}$ (c) $\delta_{ij}\delta_{jk}$ (d) $\varepsilon_{1jk}\delta_{3j}v_k$
15. Prove Lagrange's identity 1.3.20c.
16. If **e** is a unit vector and **a** an arbitrary vector, show that
$$\mathbf{a}=(\mathbf{a}\cdot\mathbf{e})\mathbf{e}+\mathbf{e}\times(\mathbf{a}\times\mathbf{e})$$
which is another representation of Eqn. 1.1.2, where **a** can be resolved into components parallel and perpendicular to **e**.

[5] That is, "components" are used for vectors and "coordinates" are used for points

1.4 Matrices and Element Form

1.4.1 Matrix – Matrix Multiplication

In the next section, §1.5, regarding vector transformation equations, it will be necessary to multiply various matrices with each other (of sizes 3×1, 1×3 and 3×3). It will be helpful to write these matrix multiplications in a short-hand element form and to develop some short "rules" which will be beneficial right through this chapter.

First, it has been seen that the dot product of two vectors can be represented by $[\mathbf{u}^{\mathrm{T}}][\mathbf{v}]$, or $u_i v_i$. Similarly, the matrix multiplication $[\mathbf{u}][\mathbf{v}^{\mathrm{T}}]$ gives a 3×3 matrix with element form $u_i v_j$ or, in full,

$$\begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \end{bmatrix}$$

This type of matrix represents the **tensor product** of two vectors, written in symbolic notation as $\mathbf{u}\otimes\mathbf{v}$ (or simply **uv**). Tensor products will be discussed in detail in §1.8 and §1.9.

Next, the matrix multiplication

$$[\mathbf{Q}][\mathbf{u}] \equiv \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{1.4.1}$$

is a 3×1 matrix with elements $([\mathbf{Q}][\mathbf{u}])_i \equiv Q_{ij} u_j$ {▲Problem 1}. The elements of $[\mathbf{Q}][\mathbf{u}]$ are the same as those of $[\mathbf{u}^{\mathrm{T}}][\mathbf{Q}^{\mathrm{T}}]$, which in element form reads $([\mathbf{u}^{\mathrm{T}}][\mathbf{Q}^{\mathrm{T}}])_i \equiv u_j Q_{ij}$.

The expression $[\mathbf{u}][\mathbf{Q}]$ is meaningless, but $[\mathbf{u}^{\mathrm{T}}][\mathbf{Q}]$ {▲Problem 2} is a 1×3 matrix with elements $([\mathbf{u}^{\mathrm{T}}][\mathbf{Q}])_i \equiv u_j Q_{ji}$.

This leads to the following rule:

1. if a vector pre-multiplies a matrix $[\mathbf{Q}]$ $\rightarrow$ it is the transpose $[\mathbf{u}^{\mathrm{T}}]$
2. if a matrix $[\mathbf{Q}]$ pre-multiplies the vector $\rightarrow$ it is $[\mathbf{u}]$
3. if summed indices are "beside each other", as the *j* in $u_j Q_{ji}$ or $Q_{ij} u_j$
   $\rightarrow$ the matrix is $[\mathbf{Q}]$
4. if summed indices are not beside each other, as the *j* in $u_j Q_{ij}$
   $\rightarrow$ the matrix is the transpose, $[\mathbf{Q}^{\mathrm{T}}]$

Finally, consider the multiplication of 3×3 matrices. Again, this follows the "beside each other" rule for the summed index. For example, $[\mathbf{A}][\mathbf{B}]$ gives the 3×3 matrix {▲Problem 6} $([\mathbf{A}][\mathbf{B}])_{ij} = A_{ik}B_{kj}$, and the multiplication $[\mathbf{A}^{\mathrm{T}}][\mathbf{B}]$ is written as $([\mathbf{A}^{\mathrm{T}}][\mathbf{B}])_{ij} = A_{ki}B_{kj}$. There is also the important identity

$$([\mathbf{A}][\mathbf{B}])^{\mathrm{T}} = [\mathbf{B}^{\mathrm{T}}][\mathbf{A}^{\mathrm{T}}] \qquad (1.4.2)$$

Note also the following (which applies to both the index notation and element form):

(i) if there is no free index, as in $u_i v_i$, there is one element (representing a scalar)

(ii) if there is one free index, as in $u_j Q_{ji}$, it is a 3×1 (or 1×3) matrix (representing a vector)

(iii) if there are two free indices, as in $A_{ki}B_{kj}$, it is a 3×3 matrix (representing, as will be seen later, a second-order tensor)

1.4.2 The Trace of a Matrix

Another important notation involving matrices is the **trace** of a matrix, defined to be the sum of the diagonal terms, and denoted by

$$\boxed{\mathrm{tr}[\mathbf{A}] = A_{11} + A_{22} + A_{33} \equiv A_{ii}} \qquad \textbf{The Trace} \qquad (1.4.3)$$

1.4.3 Problems

1. Show that $([\mathbf{Q}][\mathbf{u}])_i \equiv Q_{ij}u_j$. To do this, multiply the matrix and the vector in Eqn. 1.4.1 and write out the resulting vector in full; Show that the three elements of the vector are $Q_{1j}u_j$, $Q_{2j}u_j$ and $Q_{3j}u_j$.
2. Show that $[\mathbf{u}^{\mathrm{T}}][\mathbf{Q}]$ is a 1×3 matrix with elements $u_j Q_{ji}$ (write the matrices out in full).
3. Show that $([\mathbf{Q}][\mathbf{u}])^{\mathrm{T}} = [\mathbf{u}^{\mathrm{T}}][\mathbf{Q}^{\mathrm{T}}]$.
4. Are the three elements of $[\mathbf{Q}][\mathbf{u}]$ the same as those of $[\mathbf{u}^{\mathrm{T}}][\mathbf{Q}]$?
5. What is the element form for the matrix representation of $(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$?
6. Write out the 3×3 matrices **A** and **B** in full, i.e. in terms of A_{11}, A_{12}, etc. and verify that $[\mathbf{AB}]_{ij} = A_{ik}B_{kj}$ for $i = 2,\ j = 1$.
7. What is the element form for
   (i) $[\mathbf{A}][\mathbf{B}^{\mathrm{T}}]$
   (ii) $[\mathbf{v}^{\mathrm{T}}][\mathbf{A}][\mathbf{v}]$ (there is no ambiguity here, since $([\mathbf{v}^{\mathrm{T}}][\mathbf{A}])[\mathbf{v}] = [\mathbf{v}^{\mathrm{T}}]([\mathbf{A}][\mathbf{v}])$)
   (iii) $[\mathbf{B}^{\mathrm{T}}][\mathbf{A}][\mathbf{B}]$
8. Show that $\delta_{ij}A_{ij} = \mathrm{tr}[\mathbf{A}]$.
9. Show that $\det[\mathbf{A}] = \varepsilon_{ijk}A_{1i}A_{2j}A_{3k} = \varepsilon_{ijk}A_{i1}A_{j2}A_{k3}$.

1.5 Coordinate Transformation of Vector Components

Very often in practical problems, the components of a vector are known in one coordinate system but it is necessary to find them in some other coordinate system.

For example, one might know that the force **f** acting "in the x_1 direction" has a certain value, Fig. 1.5.1 – this is equivalent to knowing the x_1 component of the force, in an $x_1 - x_2$ coordinate system. One might then want to know what force is "acting" in some other direction – for example in the x_1' direction shown – this is equivalent to asking what the x_1' component of the force is in a new $x_1' - x_2'$ coordinate system.

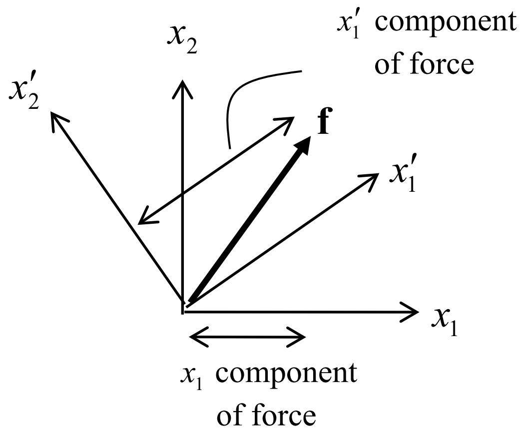


Figure 1.5.1: a vector represented using two different coordinate systems

The relationship between the components in one coordinate system and the components in a second coordinate system are called the **transformation equations**. These transformation equations are derived and discussed in what follows.

1.5.1 Rotations and Translations

Any change of Cartesian coordinate system will be due to a **translation** of the base vectors and a **rotation** of the base vectors. A translation of the base vectors does not change the components of a vector. Mathematically, this can be expressed by saying that the components of a vector **a** are $\mathbf{e}_i \cdot \mathbf{a}$, and these three quantities do not change under a translation of base vectors. Rotation of the base vectors is thus what one is concerned with in what follows.

1.5.2 Components of a Vector in Different Systems

Vectors are mathematical objects which exist *independently of any coordinate system*. Introducing a coordinate system for the purpose of analysis, one could choose, for example, a certain Cartesian coordinate system with base vectors $\mathbf{e}_i$ and origin *o*, Fig.

1.5.2. In that case the vector can be written as $\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3$, and u_1, u_2, u_3 are its components.

Now a second coordinate system can be introduced (with the same origin), this time with base vectors $\mathbf{e}'_i$. In that case, the vector can be written as $\mathbf{u} = u'_1\mathbf{e}'_1 + u'_2\mathbf{e}'_2 + u'_3\mathbf{e}'_3$, where u'_1, u'_2, u'_3 are its components in this second coordinate system, as shown in the figure. Thus the *same vector* can be written in more than one way:

$$\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3 = u'_1\mathbf{e}'_1 + u'_2\mathbf{e}'_2 + u'_3\mathbf{e}'_3 \tag{1.5.1}$$

The first coordinate system is often referred to as "the $ox_1x_2x_3$ system" and the second as "the $ox'_1x'_2x'_3$ system".

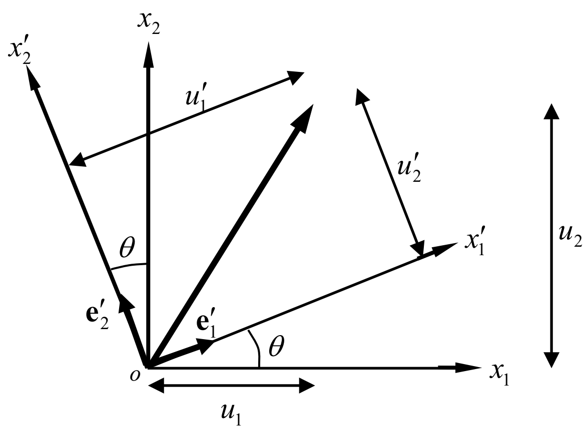


Figure 1.5.2: a vector represented using two different coordinate systems

Note that the new coordinate system is obtained from the first one by a *rotation* of the base vectors. The figure shows a rotation θ about the x_3 axis (the sign convention for rotations is positive counterclockwise).

Two Dimensions

Concentrating for the moment on the two dimensions $x_1 - x_2$, from trigonometry (refer to Fig. 1.5.3),

$$\begin{aligned}\mathbf{u} &= u_1\mathbf{e}_1 + u_2\mathbf{e}_2 \\ &= \left[|OB| - |AB|\right]\mathbf{e}_1 + \left[|BD| + |CP|\right]\mathbf{e}_2 \\ &= \left[\cos\theta\, u'_1 - \sin\theta\, u'_2\right]\mathbf{e}_1 + \left[\sin\theta\, u'_1 + \cos\theta\, u'_2\right]\mathbf{e}_2\end{aligned} \tag{1.5.2}$$

and so

$$\begin{aligned} u_1 &= \cos\theta\, u_1' - \sin\theta\, u_2' \\ u_2 &= \sin\theta\, u_1' + \cos\theta\, u_2' \end{aligned}$$

vector components in first coordinate system

vector components in second coordinate system

In matrix form, these transformation equations can be written as

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} \tag{1.5.3}$$

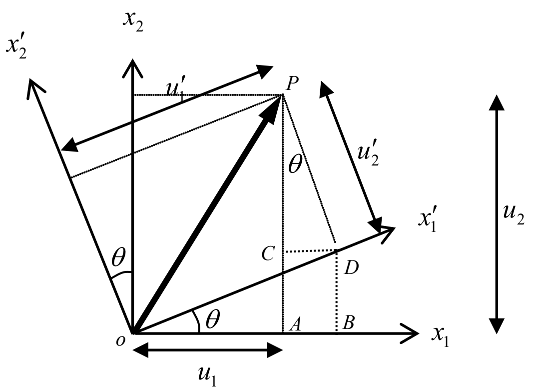


Figure 1.5.3: geometry of the 2D coordinate transformation

The 2×2 matrix is called the **transformation** or **rotation matrix** $[\mathbf{Q}]$. By pre-multiplying both sides of these equations by the inverse of $[\mathbf{Q}]$, $[\mathbf{Q}^{-1}]$, one obtains the transformation equations transforming from $[u_1 \quad u_2]^{\mathrm{T}}$ to $[u_1' \quad u_2']^{\mathrm{T}}$:

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{1.5.4}$$

An important property of the transformation matrix is that it is **orthogonal**, by which is meant that

$$\boxed{[\mathbf{Q}^{-1}] = [\mathbf{Q}^{\mathrm{T}}]} \qquad \textbf{Orthogonality of Transformation/Rotation Matrix} \tag{1.5.5}$$

Three Dimensions

The three dimensional case is shown in Fig. 1.5.4a. In this more general case, note that

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \cdot \mathbf{u} \\ \mathbf{e}_2 \cdot \mathbf{u} \\ \mathbf{e}_3 \cdot \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \cdot \left(u_1'\mathbf{e}_1' + u_2'\mathbf{e}_2' + u_3'\mathbf{e}_3'\right) \\ \mathbf{e}_2 \cdot \left(u_1'\mathbf{e}_1' + u_2'\mathbf{e}_2' + u_3'\mathbf{e}_3'\right) \\ \mathbf{e}_3 \cdot \left(u_1'\mathbf{e}_1' + u_2'\mathbf{e}_2' + u_3'\mathbf{e}_3'\right) \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \cdot \mathbf{e}_1' & \mathbf{e}_1 \cdot \mathbf{e}_2' & \mathbf{e}_1 \cdot \mathbf{e}_3' \\ \mathbf{e}_2 \cdot \mathbf{e}_1' & \mathbf{e}_2 \cdot \mathbf{e}_2' & \mathbf{e}_2 \cdot \mathbf{e}_3' \\ \mathbf{e}_3 \cdot \mathbf{e}_1' & \mathbf{e}_3 \cdot \mathbf{e}_2' & \mathbf{e}_3 \cdot \mathbf{e}_3' \end{bmatrix} \begin{bmatrix} u_1' \\ u_2' \\ u_3' \end{bmatrix} \tag{1.5.6}$$

The dot products of the base vectors from the two different coordinate systems can be seen to be the cosines of the angles between the coordinate axes. This is illustrated in Fig. 1.5.4b for the case of $\mathbf{e}_1' \cdot \mathbf{e}_j$. In general:

$$\mathbf{e}_i \cdot \mathbf{e}_j' = \cos\left(x_i, x_j'\right) \tag{1.5.7}$$

The nine quantities $\cos(x_i, x_j')$ are called the **direction cosines**, and Eqn. 1.5.6 can be expressed alternatively as

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \cos(x_1, x_1') & \cos(x_1, x_2') & \cos(x_1, x_3') \\ \cos(x_2, x_1') & \cos(x_2, x_2') & \cos(x_2, x_3') \\ \cos(x_3, x_1') & \cos(x_3, x_2') & \cos(x_3, x_3') \end{bmatrix} \begin{bmatrix} u_1' \\ u_2' \\ u_3' \end{bmatrix} \tag{1.5.8}$$

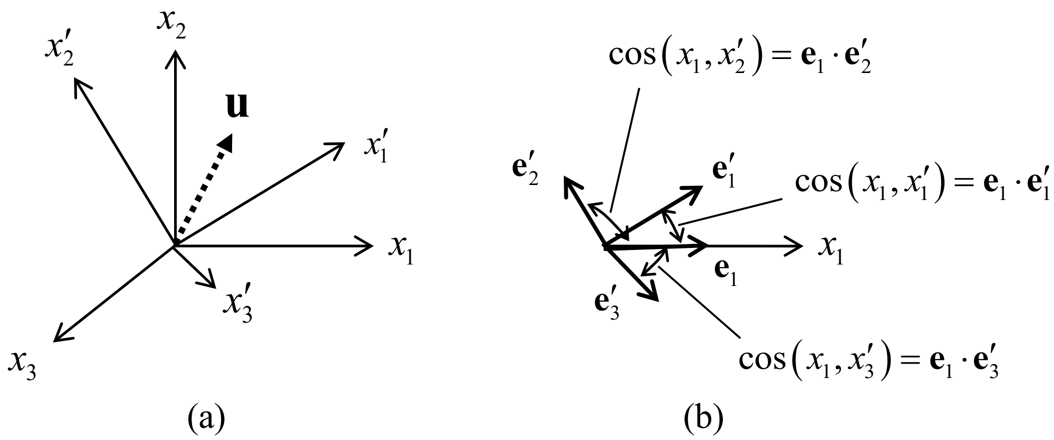


Figure 1.5.4: a 3D space: (a) two different coordinate systems, (b) direction cosines

Again denoting the components of this transformation matrix by the letter Q, $Q_{11} = \cos(x_1, x_1'), Q_{12} = \cos(x_1, x_2')$, etc., so that

$$Q_{ij} = \cos(x_i, x_j') = \mathbf{e}_i \cdot \mathbf{e}_j' . \tag{1.5.9}$$

One has the general 3D transformation matrix equations

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \begin{bmatrix} u_1' \\ u_2' \\ u_3' \end{bmatrix} \tag{1.5.10}$$

or, in element form and short-hand matrix notation,

$$u_i = Q_{ij} u'_j \qquad \ldots \quad [\mathbf{u}] = [\mathbf{Q}][\mathbf{u}'] \tag{1.5.11}$$

Note: some authors define the matrix of direction cosines to consist of the components $Q_{ij} = \cos(x'_i, x_j)$, so that the subscript *i* refers to the new coordinate system and the *j* to the old coordinate system, rather than the other way around as used here.

Formal Derivation of the Transformation Equations

The above derivation of the transformation equations Eqns. 1.5.11, $u_i = Q_{ij} u'_j$, is here carried out again using the index notation in a concise manner: start with the relations $\mathbf{u} = u_k \mathbf{e}_k = u'_j \mathbf{e}'_j$ and post-multiply both sides by $\mathbf{e}_i$ to get (the corresponding matrix representation is to the right (also, see Problem 3 in §1.4.3)):

$$\begin{aligned} & u_k \mathbf{e}_k \cdot \mathbf{e}_i = u'_j \mathbf{e}'_j \cdot \mathbf{e}_i \\ & \rightarrow u_k \delta_{ki} = u'_j Q_{ij} \\ & \rightarrow \quad u_i = u'_j Q_{ij} \qquad \ldots \quad \left[\mathbf{u}^{\mathrm{T}}\right] = \left[\mathbf{u}'^{\mathrm{T}}\right]\left[\mathbf{Q}^{\mathrm{T}}\right] \\ & \rightarrow \quad u_i = Q_{ij} u'_j \qquad \ldots \quad [\mathbf{u}] = [\mathbf{Q}][\mathbf{u}'] \end{aligned} \tag{1.5.12}$$

The inverse equations are {▲Problem 3}

$$u'_i = Q_{ji} u_j \qquad \ldots \quad [\mathbf{u}'] = \left[\mathbf{Q}^{\mathrm{T}}\right][\mathbf{u}] \tag{1.5.13}$$

Orthogonality of the Transformation Matrix $[\mathbf{Q}]$

As in the two dimensional case, the transformation matrix is orthogonal, $\left[\mathbf{Q}^{\mathrm{T}}\right] = \left[\mathbf{Q}^{-1}\right]$. This follows from 1.5.11, 1.5.13.

Example

Consider a Cartesian coordinate system with base vectors $\mathbf{e}_i$. A coordinate transformation is carried out with the new basis given by

$$\begin{aligned} \mathbf{e}'_1 &= n_1^{(1)} \mathbf{e}_1 + n_2^{(1)} \mathbf{e}_2 + n_3^{(1)} \mathbf{e}_3 \\ \mathbf{e}'_2 &= n_1^{(2)} \mathbf{e}_1 + n_2^{(2)} \mathbf{e}_2 + n_3^{(2)} \mathbf{e}_3 \\ \mathbf{e}'_3 &= n_1^{(3)} \mathbf{e}_1 + n_2^{(3)} \mathbf{e}_2 + n_3^{(3)} \mathbf{e}_3 \end{aligned}$$

What is the transformation matrix?

<u>Solution</u>

The transformation matrix consists of the direction cosines $Q_{ij} = \cos(x_i, x'_j) = \mathbf{e}_i \cdot \mathbf{e}'_j$, so

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} n_1^{(1)} & n_1^{(2)} & n_1^{(3)} \\ n_2^{(1)} & n_2^{(2)} & n_2^{(3)} \\ n_3^{(1)} & n_3^{(2)} & n_3^{(3)} \end{bmatrix} \begin{bmatrix} u_1' \\ u_2' \\ u_3' \end{bmatrix}$$

■

1.5.3 Problems

1. The angles between the axes in two coordinate systems are given in the table below.

	x_1	x_2	x_3
x_1'	135°	60°	120°
x_2'	90°	45°	45°
x_3'	45°	60°	120°

   Construct the corresponding transformation matrix $[\mathbf{Q}]$ and verify that it is orthogonal.

2. The $ox_1'x_2'x_3'$ coordinate system is obtained from the $ox_1x_2x_3$ coordinate system by a positive (counterclockwise) rotation of θ about the x_3 axis. Find the (full three dimensional) transformation matrix $[\mathbf{Q}]$. A further positive rotation β about the x_2 axis is then made to give the $ox_1''x_2''x_3''$ coordinate system. Find the corresponding transformation matrix $[\mathbf{P}]$. Then construct the transformation matrix $[\mathbf{R}]$ for the complete transformation from the $ox_1x_2x_3$ to the $ox_1''x_2''x_3''$ coordinate system.

3. Beginning with the expression $u_j\mathbf{e}_j \cdot \mathbf{e}_i' = u_k'\mathbf{e}_k' \cdot \mathbf{e}_i'$, formally derive the relation $u_i' = Q_{ji}u_j$ ($[\mathbf{u}'] = [\mathbf{Q}^{\mathrm{T}}][\mathbf{u}]$).

1.6 Vector Calculus 1 - Differentiation

Calculus involving vectors is discussed in this section, rather intuitively at first and more formally toward the end of this section.

1.6.1 The Ordinary Calculus

Consider a **scalar-valued function of a scalar**, for example the time-dependent density of a material $\rho = \rho(t)$. The calculus of scalar valued functions of scalars is just the ordinary calculus. Some of the important concepts of the ordinary calculus are reviewed in Appendix B to this Chapter, §1.B.2.

1.6.2 Vector-valued Functions of a scalar

Consider a **vector-valued function of a scalar**, for example the time-dependent displacement of a particle $\mathbf{u} = \mathbf{u}(t)$. In this case, the derivative is defined in the usual way,

$$\frac{d\mathbf{u}}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{u}(t+\Delta t) - \mathbf{u}(t)}{\Delta t},$$

which turns out to be simply the derivative of the coefficients[1],

$$\frac{d\mathbf{u}}{dt} = \frac{du_1}{dt}\mathbf{e}_1 + \frac{du_2}{dt}\mathbf{e}_2 + \frac{du_3}{dt}\mathbf{e}_3 \equiv \frac{du_i}{dt}\mathbf{e}_i$$

Partial derivatives can also be defined in the usual way. For example, if **u** is a function of the coordinates, $\mathbf{u}(x_1, x_2, x_3)$, then

$$\frac{\partial \mathbf{u}}{\partial x_1} = \lim_{\Delta x_1 \to 0} \frac{\mathbf{u}(x_1 + \Delta x_1, x_2, x_3) - \mathbf{u}(x_1, x_2, x_3)}{\Delta x_1}$$

Differentials of vectors are also defined in the usual way, so that when u_1, u_2, u_3 undergo increments $du_1 = \Delta u_1$, $du_2 = \Delta u_2$, $du_3 = \Delta u_3$, the differential of **u** is

$$d\mathbf{u} = du_1\mathbf{e}_1 + du_2\mathbf{e}_2 + du_3\mathbf{e}_3$$

and the differential and actual increment $\Delta\mathbf{u}$ approach one another as $\Delta u_1, \Delta u_2, \Delta u_3 \to 0$.

[1] assuming that the base vectors do not depend on *t*

Space Curves

The derivative of a vector can be interpreted geometrically as shown in Fig. 1.6.1: $\Delta\mathbf{u}$ is the increment in $\mathbf{u}$ consequent upon an increment Δt in t. As t changes, the end-point of the vector $\mathbf{u}(t)$ traces out the dotted curve Γ shown – it is clear that as $\Delta t \to 0$, $\Delta\mathbf{u}$ approaches the tangent to Γ, so that $d\mathbf{u}/dt$ is tangential to Γ. The unit vector tangent to the curve is denoted by $\boldsymbol{\tau}$:

$$\boldsymbol{\tau} = \frac{d\mathbf{u}/dt}{|d\mathbf{u}/dt|} \tag{1.6.1}$$

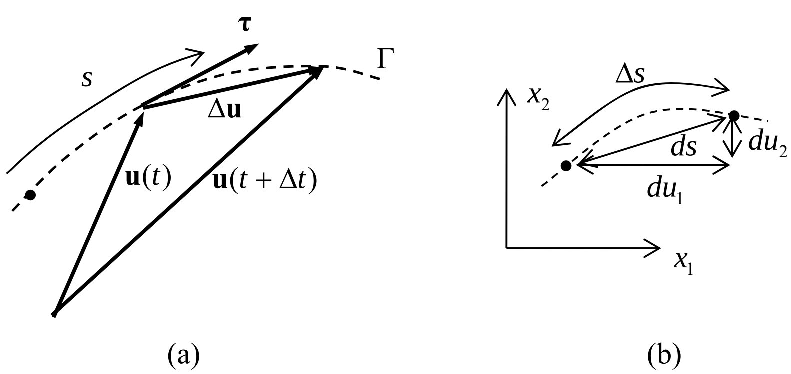


Figure 1.6.1: a space curve; (a) the tangent vector, (b) increment in arc length

Let s be a measure of the length of the curve Γ, measured from some fixed point on Γ. Let Δs be the increment in arc-length corresponding to increments in the coordinates, $\Delta\mathbf{u} = [\Delta u_1, \Delta u_2, \Delta u_3]^{\mathrm{T}}$, Fig. 1.6.1b. Then, from the ordinary calculus (see Appendix 1.B),

$$(ds)^2 = (du_1)^2 + (du_2)^2 + (du_3)^2$$

so that

$$\frac{ds}{dt} = \sqrt{\left(\frac{du_1}{dt}\right)^2 + \left(\frac{du_2}{dt}\right)^2 + \left(\frac{du_3}{dt}\right)^2}$$

But

$$\frac{d\mathbf{u}}{dt} = \frac{du_1}{dt}\mathbf{e}_1 + \frac{du_2}{dt}\mathbf{e}_2 + \frac{du_3}{dt}\mathbf{e}_3$$

so that

$$\left|\frac{d\mathbf{u}}{dt}\right| = \frac{ds}{dt} \tag{1.6.2}$$

Thus the unit vector tangent to the curve can be written as

$$\boldsymbol{\tau} = \frac{d\mathbf{u}/dt}{ds/dt} = \frac{d\mathbf{u}}{ds} \tag{1.6.3}$$

If $\mathbf{u}$ is interpreted as the position vector of a particle and t is interpreted as time, then $\mathbf{v} = d\mathbf{u}/dt$ is the velocity vector of the particle as it moves with speed ds/dt along Γ.

Example (of particle motion)

A particle moves along a curve whose parametric equations are $x_1 = 2t^2$, $x_2 = t^2 - 4t$, $x_3 = 3t - 5$ where t is time. Find the component of the velocity at time $t = 1$ in the direction $\mathbf{a} = \mathbf{e}_1 - 3\mathbf{e}_2 + 2\mathbf{e}_3$.

<u>Solution</u>

The velocity is

$$\begin{aligned}\mathbf{v} &= \frac{d\mathbf{r}}{dt} = \frac{d}{dt}\left\{2t^2\mathbf{e}_1 + \left(t^2 - 4t\right)\mathbf{e}_2 + \left(3t - 5\right)\mathbf{e}_3\right\} \\ &= 4\mathbf{e}_1 - 2\mathbf{e}_2 + 3\mathbf{e}_3 \quad \text{at } t = 1\end{aligned}$$

The component in the given direction is $\mathbf{v} \cdot \hat{\mathbf{a}}$, where $\hat{\mathbf{a}}$ is a unit vector in the direction of $\mathbf{a}$, giving $8\sqrt{14}/7$.

■

Curvature

The scalar **curvature** $\kappa(s)$ of a space curve is defined to be the magnitude of the rate of change of the unit tangent vector:

$$\kappa(s) = \left|\frac{d\boldsymbol{\tau}}{ds}\right| = \left|\frac{d^2\mathbf{u}}{ds^2}\right| \tag{1.6.4}$$

Note that $\Delta\boldsymbol{\tau}$ is in a direction perpendicular to $\boldsymbol{\tau}$, Fig. 1.6.2. In fact, this can be proved as follows: since $\boldsymbol{\tau}$ is a unit vector, $\boldsymbol{\tau} \cdot \boldsymbol{\tau}$ is a constant ($=1$), and so $d\left(\boldsymbol{\tau} \cdot \boldsymbol{\tau}\right)/ds = 0$, but also,

$$\frac{d}{ds}\left(\boldsymbol{\tau} \cdot \boldsymbol{\tau}\right) = 2\boldsymbol{\tau} \cdot \frac{d\boldsymbol{\tau}}{ds}$$

and so $\boldsymbol{\tau}$ and $d\boldsymbol{\tau}/ds$ are perpendicular. The unit vector defined in this way is called the **principal normal vector**:

$$\mathbf{v} = \frac{1}{\kappa}\frac{d\boldsymbol{\tau}}{ds} \tag{1.6.5}$$

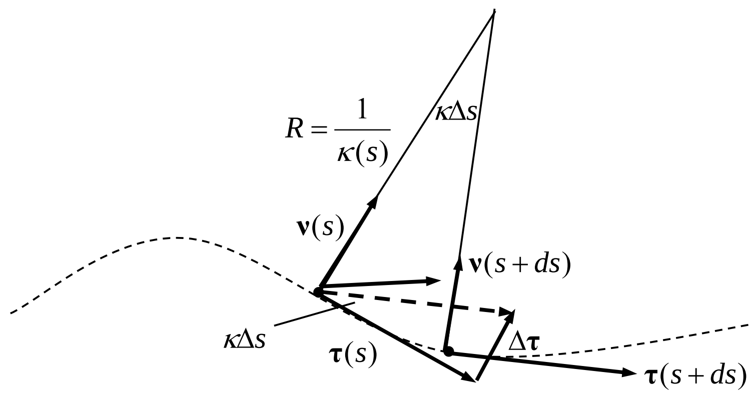


Figure 1.6.2: the curvature

This can be seen geometrically in Fig. 1.6.2: from Eqn. 1.6.5, $\Delta\boldsymbol{\tau}$ is a vector of magnitude $\kappa\Delta s$ in the direction of the vector normal to $\boldsymbol{\tau}$. The **radius of curvature** *R* is defined as the reciprocal of the curvature; it is the radius of the circle which just touches the curve at *s*, Fig. 1.6.2.

Finally, the unit vector perpendicular to both the tangent vector and the principal normal vector is called the **unit binormal vector**:

$$\mathbf{b} = \boldsymbol{\tau} \times \mathbf{v} \tag{1.6.6}$$

The planes defined by these vectors are shown in Fig. 1.6.3; they are called the **rectifying plane**, the **normal plane** and the **osculating plane**.

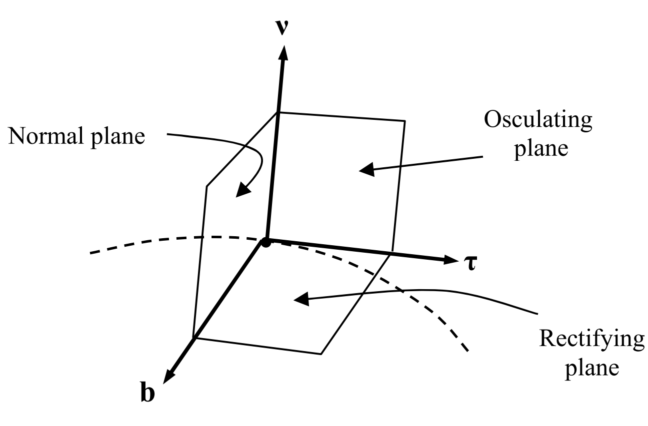


Figure 1.6.3: the unit tangent, principal normal and binormal vectors and associated planes

Rules of Differentiation

The derivative of a vector is also a vector and the usual rules of differentiation apply,

$$\begin{aligned}\frac{d}{dt}(\mathbf{u}+\mathbf{v}) &= \frac{d\mathbf{u}}{dt}+\frac{d\mathbf{v}}{dt}\\ \frac{d}{dt}(\alpha(t)\mathbf{v}) &= \alpha\frac{d\mathbf{v}}{dt}+\mathbf{v}\frac{d\alpha}{dt}\end{aligned} \tag{1.6.7}$$

Also, it is straight forward to show that {▲Problem 2}

$$\frac{d}{dt}(\mathbf{v}\cdot\mathbf{a}) = \mathbf{v}\cdot\frac{d\mathbf{a}}{dt}+\frac{d\mathbf{v}}{dt}\cdot\mathbf{a} \qquad \frac{d}{dt}(\mathbf{v}\times\mathbf{a}) = \mathbf{v}\times\frac{d\mathbf{a}}{dt}+\frac{d\mathbf{v}}{dt}\times\mathbf{a} \tag{1.6.8}$$

(The order of the terms in the cross-product expression is important here.)

1.6.3 Fields

In many applications of vector calculus, a scalar or vector can be associated with each point in space **x**. In this case they are called **scalar** or **vector fields**. For example

$\theta(\mathbf{x})$	temperature	a scalar field (a scalar-valued function of position)
$\mathbf{v}(\mathbf{x})$	velocity	a vector field (a vector valued function of position)

These quantities will in general depend also on time, so that one writes $\theta(\mathbf{x},t)$ or $\mathbf{v}(\mathbf{x},t)$. Partial differentiation of scalar and vector fields with respect to the variable *t* is symbolised by $\partial/\partial t$. On the other hand, partial differentiation with respect to the coordinates is symbolised by $\partial/\partial x_i$. The notation can be made more compact by introducing the **subscript comma** to denote partial differentiation with respect to the coordinate variables, in which case $\phi_{,i} = \partial\phi/\partial x_i$, $u_{i,jk} = \partial^2 u_i/\partial x_j \partial x_k$, and so on.

1.6.4 The Gradient of a Scalar Field

Let $\phi(\mathbf{x})$ be a scalar field. The **gradient** of ϕ is a vector field defined by (see Fig. 1.6.4)

$$\boxed{\begin{aligned}\nabla\phi &= \frac{\partial\phi}{\partial x_1}\mathbf{e}_1+\frac{\partial\phi}{\partial x_2}\mathbf{e}_2+\frac{\partial\phi}{\partial x_3}\mathbf{e}_3\\ &= \frac{\partial\phi}{\partial x_i}\mathbf{e}_i\\ &\equiv \frac{\partial\phi}{\partial\mathbf{x}}\end{aligned}} \qquad \textbf{Gradient of a Scalar Field} \tag{1.6.9}$$

The gradient $\nabla\phi$ is of considerable importance because if one takes the dot product of $\nabla\phi$ with $d\mathbf{x}$, it gives the increment in ϕ:

$$\begin{aligned}\nabla\phi\cdot d\mathbf{x} &= \frac{\partial\phi}{\partial x_i}\mathbf{e}_i\cdot dx_j\mathbf{e}_j \\ &= \frac{\partial\phi}{\partial x_i}dx_i \\ &= d\phi \\ &= \phi(\mathbf{x}+d\mathbf{x})-\phi(d\mathbf{x})\end{aligned} \tag{1.6.10}$$

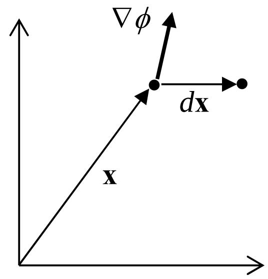


Figure 1.6.4: the gradient of a vector

If one writes $d\mathbf{x}$ as $|d\mathbf{x}|\mathbf{e} = dx\mathbf{e}$, where **e** is a unit vector in the direction of $d\mathbf{x}$, then

$$\nabla\phi\cdot\mathbf{e} = \left(\frac{d\phi}{dx}\right)_{\text{in } \mathbf{e} \text{ direction}} \equiv \frac{d\phi}{dn} \tag{1.6.11}$$

This quantity is called the **directional derivative** of ϕ, in the direction of **e**, and will be discussed further in §1.6.11.

The gradient of a scalar field is also called the **scalar gradient**, to distinguish it from the **vector gradient** (see later)[2], and is also denoted by

$$\text{grad}\,\phi \equiv \nabla\phi \tag{1.6.12}$$

Example (of the Gradient of a Scalar Field)

Consider a two-dimensional temperature field $\theta = x_1^2 + x_2^2$. Then

$$\nabla\theta = 2x_1\mathbf{e}_1 + 2x_2\mathbf{e}_2$$

For example, at $(1,0)$, $\theta = 1$, $\nabla\theta = 2\mathbf{e}_1$ and at $(1,1)$, $\theta = 2$, $\nabla\theta = 2\mathbf{e}_1 + 2\mathbf{e}_2$, Fig. 1.6.5. Note the following:

(i) $\nabla\theta$ points in the direction *normal* to the curve $\theta = \text{const.}$
(ii) the direction of *maximum* rate of change of θ is in the direction of $\nabla\theta$

[2] in this context, a *gradient* is a derivative with respect to a position vector, but the term gradient is used more generally than this, e.g. see §1.14

(iii) the direction of zero $d\theta$ is in the direction *perpendicular* to $\nabla\theta$

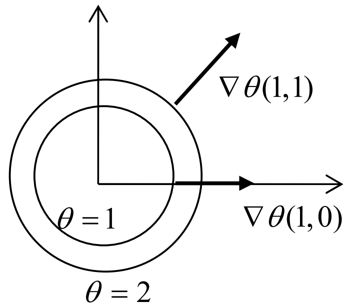


Figure 1.6.5: gradient of a temperature field

The curves $\theta(x_1, x_2) = \text{const.}$ are called **isotherms** (curves of constant temperature). In general, they are called **iso-curves** (or **iso-surfaces** in three dimensions).

■

Many physical laws are given in terms of the gradient of a scalar field. For example, **Fourier's law** of heat conduction relates the heat flux **q** (the rate at which heat flows through a surface of unit area[3]) to the temperature gradient through

$$\mathbf{q} = -k\nabla\theta \tag{1.6.13}$$

where *k* is the **thermal conductivity** of the material, so that heat flows along the direction normal to the isotherms.

The Normal to a Surface

In the above example, it was seen that $\nabla\theta$ points in the direction normal to the curve $\theta = \text{const.}$ Here it will be seen generally how and why the gradient can be used to obtain a normal vector to a surface.

Consider a surface represented by the scalar function $f(x_1, x_2, x_3) = c$, c a constant[4], and also a space curve C lying on the surface, defined by the position vector $\mathbf{r} = x_1(t)\mathbf{e}_1 + x_2(t)\mathbf{e}_2 + x_3(t)\mathbf{e}_3$. The components of **r** must satisfy the equation of the surface, so $f(x_1(t), x_2(t), x_3(t)) = c$. Differentiation gives

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial f}{\partial x_3}\frac{dx_3}{dt} = 0$$

---

[3] the **flux** is the rate of flow of fluid, particles or energy through a given surface; the **flux density** is the flux per unit area but, as here, this is more commonly referred to simply as the flux

[4] a surface can be represented by the equation $f(x_1, x_2, x_3) = c$; for example, the expression $x_1^2 + x_2^2 + x_3^2 = 4$ is the equation for a sphere of radius 2 (with centre at the origin). Alternatively, the surface can be written in the form $x_3 = g(x_1, x_2)$, for example $x_3 = \sqrt{4 - x_1^2 - x_2^2}$

which is equivalent to the equation $\operatorname{grad} f \cdot (d\mathbf{r}/dt) = 0$ and, as seen in §1.6.2, $d\mathbf{r}/dt$ is a vector tangential to the surface. Thus $\operatorname{grad} f$ is normal to the tangent vector; $\operatorname{grad} f$ must be normal to all the tangents to all the curves through *p*, so it must be normal to the plane tangent to the surface.

Taylor's Series

Writing ϕ as a function of three variables (omitting time *t*), so that $\phi = \phi(x_1, x_2, x_3)$, then ϕ can be expanded in a three-dimensional Taylor's series:

$$\phi(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3) = \phi(x_1, x_2, x_3) + \left\{ \frac{\partial \phi}{\partial x_1} dx_1 + \frac{\partial \phi}{\partial x_2} dx_2 + \frac{\partial \phi}{\partial x_3} dx_3 \right\} + \frac{1}{2} \left\{ \frac{\partial^2 \phi}{\partial x_1^2} (dx_1)^2 + \cdots \right\}$$

Neglecting the higher order terms, this can be written as

$$\phi(\mathbf{x} + d\mathbf{x}) = \phi(\mathbf{x}) + \frac{\partial \phi}{\partial \mathbf{x}} \cdot d\mathbf{x}$$

which is equivalent to 1.6.9, 1.6.10.

1.6.5 The Nabla Operator

The symbolic vector operator ∇ is called the **Nabla operator**[5]. One can write this in component form as

$$\nabla = \mathbf{e}_1 \frac{\partial}{\partial x_1} + \mathbf{e}_2 \frac{\partial}{\partial x_2} + \mathbf{e}_3 \frac{\partial}{\partial x_3} = \mathbf{e}_i \frac{\partial}{\partial x_i} \tag{1.6.14}$$

One can generalise the idea of the gradient of a scalar field by defining the dot product and the cross product of the vector operator ∇ with a vector field $(\bullet)$, according to the rules

$$\nabla \cdot (\bullet) = \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot (\bullet), \quad \nabla \times (\bullet) = \mathbf{e}_i \frac{\partial}{\partial x_i} \times (\bullet) \tag{1.6.15}$$

The following terminology is used:

$$\begin{aligned} \operatorname{grad} \phi &= \nabla \phi \\ \operatorname{div} \mathbf{u} &= \nabla \cdot \mathbf{u} \\ \operatorname{curl} \mathbf{u} &= \nabla \times \mathbf{u} \end{aligned} \tag{1.6.16}$$

[5] or **del** or the **Gradient operator**

These latter two are discussed in the following sections.

1.6.6 The Divergence of a Vector Field

From the definition (1.6.15), the **divergence** of a vector field $\mathbf{a}(\mathbf{x})$ is the scalar field

$$\boxed{\begin{aligned}\operatorname{div}\mathbf{a} = \nabla\cdot\mathbf{a} &= \left(\mathbf{e}_i \frac{\partial}{\partial x_i}\right)\cdot\left(a_j\mathbf{e}_j\right) = \frac{\partial a_i}{\partial x_i} \\ &= \frac{\partial a_1}{\partial x_1} + \frac{\partial a_2}{\partial x_2} + \frac{\partial a_3}{\partial x_3}\end{aligned}}$$

Divergence of a Vector Field (1.6.17)

Differential Elements & Physical interpretations of the Divergence

Consider a flowing compressible[6] material with velocity field $\mathbf{v}(x_1, x_2, x_3)$. Consider now a **differential element** of this material, with dimensions $\Delta x_1, \Delta x_2, \Delta x_3$, with bottom left-hand corner at (x_1, x_2, x_3), fixed in space and through which the material flows[7], Fig. 1.6.6.

The component of the velocity in the x_1 direction, v_1, will vary over a face of the element but, *if the element is small*, the velocities will vary linearly as shown; only the components at the four corners of the face are shown for clarity.

Since [distance = time $\times$ velocity], the volume of material flowing through the right-hand face in time Δt is Δt times the "volume" bounded by the four corner velocities (between the right-hand face and the plane surface denoted by the dotted lines); it is straightforward to show that this volume is equal to the volume shown to the right, Fig. 1.6.6b, with constant velocity equal to the average velocity v_{ave}, which occurs at the centre of the face. Thus the volume of material flowing out is[8] $\Delta x_2 \Delta x_3 v_{ave} \Delta t$ and the **volume flux**, i.e. the *rate* of volume flow, is $\Delta x_2 \Delta x_3 v_{ave}$. Now

$$v_{ave} = v_1(x_1 + \Delta x_1, x_2 + \tfrac{1}{2}\Delta x_2, x_3 + \tfrac{1}{2}\Delta x_3)$$

Using a Taylor's series expansion, and neglecting higher order terms,

$$v_{ave} \approx v_1(x_1, x_2, x_3) + \Delta x_1 \frac{\partial v_1}{\partial x_1} + \tfrac{1}{2}\Delta x_2 \frac{\partial v_1}{\partial x_2} + \tfrac{1}{2}\Delta x_3 \frac{\partial v_1}{\partial x_3}$$

---

[6] that is, it can be compressed or expanded

[7] this type of fixed volume in space, used in analysis, is called a **control volume**

[8] the velocity will change by a small amount during the time interval Δt. One could use the average velocity in the calculation, i.e. $\frac{1}{2}\left(v_1(\mathbf{x},t) + v_1(\mathbf{x}, t+\Delta t)\right)$, but in the limit as $\Delta t \to 0$, this will reduce to $v_1(\mathbf{x},t)$

with the partial derivatives evaluated at (x_1, x_2, x_3), so the volume flux out is

$$\Delta x_2 \Delta x_3 \left\{ v_1(x_1, x_2, x_3) + \Delta x_1 \frac{\partial v_1}{\partial x_1} + \tfrac{1}{2}\Delta x_2 \frac{\partial v_1}{\partial x_2} + \tfrac{1}{2}\Delta x_3 \frac{\partial v_1}{\partial x_3} \right\}$$

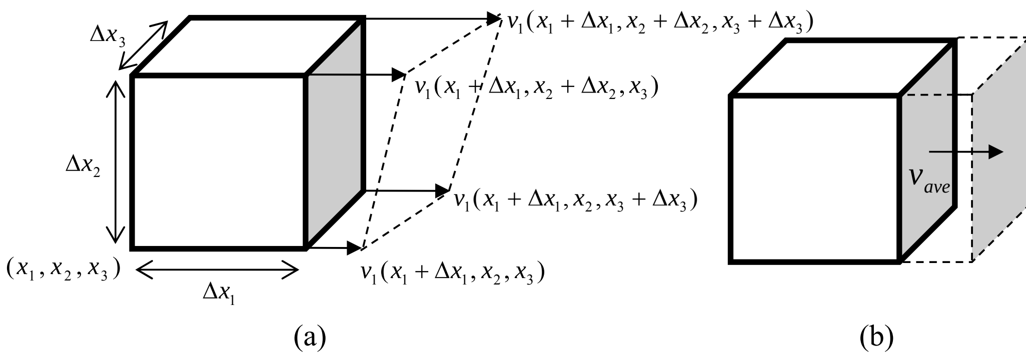


Figure 1.6.6: a differential element; (a) flow through a face, (b) volume of material flowing through the face

The net volume flux out (rate of volume flow out through the right-hand face minus the rate of volume flow in through the left-hand face) is then $\Delta x_1 \Delta x_2 \Delta x_3 (\partial v_1 / \partial x_1)$ and the net volume flux per unit volume is $\partial v_1 / \partial x_1$. Carrying out a similar calculation for the other two coordinate directions leads to

$$\textit{net unit volume flux out of an elemental volume}: \quad \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3} \equiv \operatorname{div} \mathbf{v} \qquad (1.6.18)$$

which is the physical meaning of the divergence of the velocity field.

If $\operatorname{div}\mathbf{v} > 0$, there is a net flow out and the density of material is decreasing. On the other hand, if $\operatorname{div}\mathbf{v} = 0$, the inflow equals the outflow and the density remains constant – such a material is called **incompressible**[9]. A flow which is divergence free is said to be **isochoric**. A vector $\mathbf{v}$ for which $\operatorname{div}\mathbf{v} = 0$ is said to be **solenoidal**.

<u>Notes:</u>
- The above result holds only in the limit when the element shrinks to zero size – so that the extra terms in the Taylor series tend to zero and the velocity field varies in a linear fashion over a face
- consider the velocity at a fixed point in space, $\mathbf{v}(\mathbf{x}, t)$. The velocity at a later time, $\mathbf{v}(\mathbf{x}, t + \Delta t)$, actually gives the velocity of a different material particle. This is shown in Fig. 1.6.7 below: the material particles $1, 2, 3$ are moving through space and whereas $\mathbf{v}(\mathbf{x}, t)$ represents the velocity of particle 2, $\mathbf{v}(\mathbf{x}, t + \Delta t)$ now represents the velocity of particle 1, which has moved into position $\mathbf{x}$. This point is important in the consideration of the kinematics of materials, to be discussed in Chapter 2

[9] a **liquid**, such as water, is a material which is incompressible

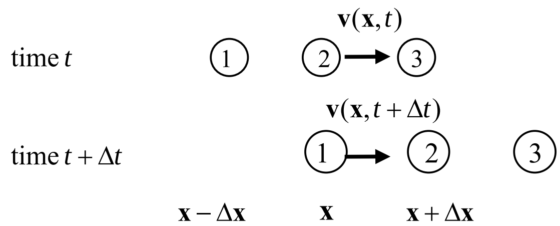


Figure 1.6.7: moving material particles

Another example would be the divergence of the heat flux vector **q**. This time suppose also that there is some generator of heat inside the element (a **source**), generating at a rate of *r* per unit volume, *r* being a scalar field. Again, assuming the element to be small, one takes *r* to be acting at the mid-point of the element, and one considers $r(x_1 + \frac{1}{2}\Delta x_1, \cdots)$. Assume a **steady-state** heat flow, so that the (heat) energy within the elemental volume remains constant with time – the law of balance of (heat) energy then requires that the net flow of heat out must equal the heat generated within, so

$$\begin{aligned}\Delta x_1 \Delta x_2 \Delta x_3 \frac{\partial q_1}{\partial x_1} + \Delta x_1 \Delta x_2 \Delta x_3 \frac{\partial q_2}{\partial x_2} + \Delta x_1 \Delta x_2 \Delta x_3 \frac{\partial q_3}{\partial x_3} \\ = \Delta x_1 \Delta x_2 \Delta x_3 \left\{ r(x_1, x_2, x_3) + \tfrac{1}{2}\Delta x_1 \frac{\partial r}{\partial x_1} + \tfrac{1}{2}\Delta x_2 \frac{\partial r}{\partial x_2} + \tfrac{1}{2}\Delta x_3 \frac{\partial r}{\partial x_3} \right\}\end{aligned}$$

Dividing through by $\Delta x_1 \Delta x_2 \Delta x_3$ and taking the limit as $\Delta x_1, \Delta x_2, \Delta x_3 \to 0$, one obtains

$$\mathrm{div}\mathbf{q} = r \tag{1.6.19}$$

Here, the divergence of the heat flux vector field can be interpreted as the heat generated (or absorbed) per unit volume per unit time in a temperature field. If the divergence is zero, there is no heat being generated (or absorbed) and the heat leaving the element is equal to the heat entering it.

1.6.7 The Laplacian

Combining Fourier's law of heat conduction (1.6.13), $\mathbf{q} = -k\nabla\theta$, with the energy balance equation (1.6.19), $\mathrm{div}\mathbf{q} = r$, and assuming the conductivity is constant, leads to $-k\,\nabla\cdot\nabla\theta = r$. Now

$$\begin{aligned}\nabla\cdot\nabla\theta &= \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot \left(\frac{\partial \theta}{\partial x_j} \mathbf{e}_j \right) = \frac{\partial}{\partial x_i}\left(\frac{\partial \theta}{\partial x_j} \right) \delta_{ij} = \frac{\partial^2 \theta}{\partial x_i^2} \\ &= \frac{\partial^2 \theta}{\partial x_1^2} + \frac{\partial^2 \theta}{\partial x_2^2} + \frac{\partial^2 \theta}{\partial x_3^2}\end{aligned} \tag{1.6.20}$$

This expression is called the **Laplacian** of θ. By introducing the Laplacian operator $\nabla^2 \equiv \nabla \cdot \nabla$, one has

$$\nabla^2 \theta = -\frac{r}{k} \tag{1.6.21}$$

This equation governs the steady state heat flow for constant conductivity. In general, the equation $\nabla^2 \phi = a$ is called **Poisson's equation**. When there are no heat sources (or sinks), one has **Laplace's equation**, $\nabla^2 \theta = 0$. Laplace's and Poisson's equation arise in many other mathematical models in mechanics, electromagnetism, etc.

1.6.8 The Curl of a Vector Field

From the definition 1.6.15 and 1.6.14, the **curl** of a vector field $\mathbf{a}(\mathbf{x})$ is the vector field

$$\boxed{\begin{aligned} \mathrm{curl}\,\mathbf{a} = \nabla \times \mathbf{a} &= \mathbf{e}_i \frac{\partial}{\partial x_i} \times \left(a_j \mathbf{e}_j\right) \\ &= \frac{\partial a_j}{\partial x_i} \mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ijk} \frac{\partial a_j}{\partial x_i} \mathbf{e}_k \end{aligned}} \quad \textbf{Curl of a Vector Field} \tag{1.6.22}$$

It can also be expressed in the form

$$\begin{aligned} \mathrm{curl}\,\mathbf{a} = \nabla \times \mathbf{a} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ a_1 & a_2 & a_3 \end{vmatrix} \\ &= \varepsilon_{ijk} \frac{\partial a_j}{\partial x_i} \mathbf{e}_k = \varepsilon_{ijk} \frac{\partial a_k}{\partial x_j} \mathbf{e}_i = \varepsilon_{ijk} \frac{\partial a_i}{\partial x_k} \mathbf{e}_j \end{aligned} \tag{1.6.23}$$

Note: the divergence and curl of a vector field are independent of any coordinate system (for example, the divergence of a vector and the length and direction of $\mathrm{curl}\,\mathbf{a}$ are independent of the coordinate system in use) – these will be re-defined without reference to any particular coordinate system when discussing tensors (see §1.14).

Physical interpretation of the Curl

Consider a particle with position vector **r** and moving with velocity $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, that is, with an angular velocity ω about an axis in the direction of $\boldsymbol{\omega}$. Then {▲Problem 7}

$$\mathrm{curl}\,\mathbf{v} = \nabla \times (\boldsymbol{\omega} \times \mathbf{r}) = 2\boldsymbol{\omega} \tag{1.6.24}$$

Thus the curl of a vector field is associated with rotational properties. In fact, if **v** is the velocity of a moving fluid, then a small paddle wheel placed in the fluid would tend to rotate in regions where $\mathrm{curl}\,\mathbf{v} \neq 0$, in which case the velocity field **v** is called a **vortex**

field. The paddle wheel would remain stationary in regions where $\mathrm{curl}\,\mathbf{v} = 0$, in which case the velocity field $\mathbf{v}$ is called **irrotational**.

1.6.9 Identities

Here are some important identities of vector calculus {▲Problem 8}:

$$
\begin{aligned}
\mathrm{grad}(\phi+\psi) &= \mathrm{grad}\phi + \mathrm{grad}\,\psi \\
\mathrm{div}(\mathbf{u}+\mathbf{v}) &= \mathrm{div}\,\mathbf{u} + \mathrm{div}\,\mathbf{v} \\
\mathrm{curl}(\mathbf{u}+\mathbf{v}) &= \mathrm{curl}\,\mathbf{u} + \mathrm{curl}\,\mathbf{v}
\end{aligned}
\tag{1.6.25}
$$

$$
\begin{aligned}
\mathrm{grad}(\phi\psi) &= \phi\,\mathrm{grad}\,\psi + \psi\,\mathrm{grad}\phi \\
\mathrm{div}(\phi\mathbf{u}) &= \phi\,\mathrm{div}\,\mathbf{u} + \mathrm{grad}\,\phi\cdot\mathbf{u} \\
\mathrm{curl}(\phi\mathbf{u}) &= \phi\,\mathrm{curl}\,\mathbf{u} + \mathrm{grad}\,\phi\times\mathbf{u} \\
\mathrm{div}(\mathbf{u}\times\mathbf{v}) &= \mathbf{v}\cdot\mathrm{curl}\,\mathbf{u} - \mathbf{u}\cdot\mathrm{curl}\,\mathbf{v} \\
\mathrm{curl}(\mathrm{grad}\phi) &= \mathbf{o} \\
\mathrm{div}(\mathrm{curl}\mathbf{u}) &= 0 \\
\mathrm{div}(\lambda\,\mathrm{grad}\phi) &= \lambda\nabla^2\phi + \mathrm{grad}\lambda\cdot\mathrm{grad}\phi
\end{aligned}
\tag{1.6.26}
$$

1.6.10 Cylindrical and Spherical Coordinates

Cartesian coordinates have been used exclusively up to this point. In many practical problems, it is easier to carry out an analysis in terms of cylindrical or spherical coordinates. Differentiation in these coordinate systems is discussed in what follows[10].

Cylindrical Coordinates

Cartesian and cylindrical coordinates are related through (see Fig. 1.6.8)

$$
\begin{aligned}
x &= r\cos\theta \qquad & r &= \sqrt{x^2+y^2} \\
y &= r\sin\theta\,, \qquad & \theta &= \tan^{-1}(y/x) \\
z &= z \qquad & z &= z
\end{aligned}
\tag{1.6.27}
$$

Then the Cartesian partial derivatives become

$$
\begin{aligned}
\frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \\
\frac{\partial}{\partial y} &= \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial y}\frac{\partial}{\partial\theta} = \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}
\end{aligned}
\tag{1.6.28}
$$

[10] this section also serves as an introduction to the more general topic of **Curvilinear Coordinates** covered in §1.16-§1.19

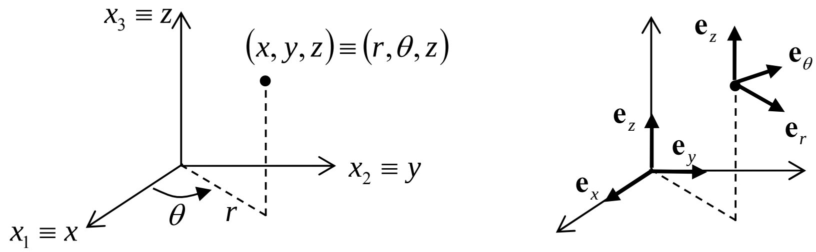


Figure 1.6.8: cylindrical coordinates

The base vectors are related through

$$\begin{aligned} \mathbf{e}_x &= \mathbf{e}_r \cos\theta - \mathbf{e}_\theta \sin\theta \qquad & \mathbf{e}_r &= \mathbf{e}_x \cos\theta + \mathbf{e}_y \sin\theta \\ \mathbf{e}_y &= \mathbf{e}_r \sin\theta + \mathbf{e}_\theta \cos\theta, \qquad & \mathbf{e}_\theta &= -\mathbf{e}_x \sin\theta + \mathbf{e}_y \cos\theta \\ \mathbf{e}_z &= \mathbf{e}_z & \mathbf{e}_z &= \mathbf{e}_z \end{aligned} \tag{1.6.29}$$

so that from Eqn. 1.6.14, after some algebra, the Nabla operator in cylindrical coordinates reads as {▲Problem 9}

$$\nabla = \mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \mathbf{e}_z \frac{\partial}{\partial z} \tag{1.6.30}$$

which allows one to take the gradient of a scalar field in cylindrical coordinates:

$$\nabla \phi = \frac{\partial \phi}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial \phi}{\partial \theta}\mathbf{e}_\theta + \frac{\partial \phi}{\partial z}\mathbf{e}_z \tag{1.6.31}$$

Cartesian base vectors are independent of position. However, the cylindrical base vectors, although they are always of unit magnitude, change direction with position. In particular, the directions of the base vectors $\mathbf{e}_r, \mathbf{e}_\theta$ depend on θ, and so these base vectors have derivatives with respect to θ: from Eqn. 1.6.29,

$$\begin{aligned} \frac{\partial}{\partial \theta}\mathbf{e}_r &= \mathbf{e}_\theta \\ \frac{\partial}{\partial \theta}\mathbf{e}_\theta &= -\mathbf{e}_r \end{aligned} \tag{1.6.32}$$

with all other derivatives of the base vectors with respect to r, θ, z equal to zero.

The divergence can now be evaluated:

$$\nabla \cdot \mathbf{v} = \left(\mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \mathbf{e}_z \frac{\partial}{\partial z} \right) \cdot \left(v_r \mathbf{e}_r + v_\theta \mathbf{e}_\theta + v_z \mathbf{e}_z \right)$$
$$= \frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z} \tag{1.6.33}$$

Similarly the curl of a vector and the Laplacian of a scalar are {▲Problem 10}

$$\nabla \times \mathbf{v} = \left(\frac{1}{r}\frac{\partial v_z}{\partial \theta} - \frac{\partial v_\theta}{\partial z} \right)\mathbf{e}_r + \left(\frac{\partial v_r}{\partial z} - \frac{\partial v_z}{\partial r} \right)\mathbf{e}_\theta + \left[\frac{1}{r}\left(\frac{\partial}{\partial r}(r v_\theta) - \frac{\partial v_r}{\partial \theta} \right) \right]\mathbf{e}_z$$
$$\nabla^2 \phi = \frac{\partial^2 \phi}{\partial r^2} + \frac{1}{r}\frac{\partial \phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \phi}{\partial \theta^2} + \frac{\partial^2 \phi}{\partial z^2} \tag{1.6.34}$$

Spherical Coordinates

Cartesian and spherical coordinates are related through (see Fig. 1.6.9)

$$\begin{aligned} x &= r \sin\theta \cos\phi \qquad & r &= \sqrt{x^2 + y^2 + z^2} \\ y &= r \sin\theta \sin\phi\,, \qquad & \theta &= \tan^{-1}\left(\sqrt{x^2+y^2} / z \right) \\ z &= r\cos\theta \qquad & \phi &= \tan^{-1}(y/x) \end{aligned} \tag{1.6.35}$$

and the base vectors are related through

$$\begin{aligned} \mathbf{e}_x &= \mathbf{e}_r \sin\theta\cos\phi + \mathbf{e}_\theta \cos\theta\cos\phi - \mathbf{e}_\phi \sin\phi \\ \mathbf{e}_y &= \mathbf{e}_r \sin\theta\sin\phi + \mathbf{e}_\theta \cos\theta\sin\phi + \mathbf{e}_\phi \cos\phi \\ \mathbf{e}_z &= \mathbf{e}_r \cos\theta - \mathbf{e}_\theta \sin\theta \\ \\ \mathbf{e}_r &= \mathbf{e}_x \sin\theta\cos\phi + \mathbf{e}_y \sin\theta\sin\phi + \mathbf{e}_z \cos\theta \\ \mathbf{e}_\theta &= \mathbf{e}_x \cos\theta\cos\phi + \mathbf{e}_y \cos\theta\sin\phi - \mathbf{e}_z \sin\theta \\ \mathbf{e}_\phi &= -\mathbf{e}_x \sin\phi + \mathbf{e}_y \cos\phi \end{aligned} \tag{1.6.36}$$

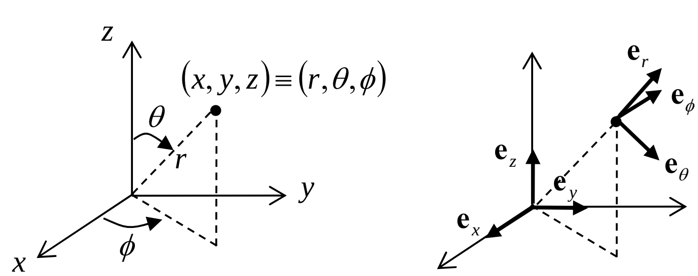


Figure 1.6.9: spherical coordinates

In this case the non-zero derivatives of the base vectors are

$$\begin{aligned}\frac{\partial}{\partial\theta}\mathbf{e}_r &= \mathbf{e}_\theta \\ \frac{\partial}{\partial\theta}\mathbf{e}_\theta &= -\mathbf{e}_r\end{aligned}, \qquad \begin{aligned}\frac{\partial}{\partial\phi}\mathbf{e}_r &= \sin\theta\mathbf{e}_\phi \\ \frac{\partial}{\partial\phi}\mathbf{e}_\theta &= \cos\theta\mathbf{e}_\phi \\ \frac{\partial}{\partial\phi}\mathbf{e}_\phi &= -\sin\theta\mathbf{e}_r - \cos\theta\mathbf{e}_\theta\end{aligned} \tag{1.6.37}$$

and it can then be shown that {▲Problem 11}

$$\begin{aligned}\nabla\varphi &= \frac{\partial\varphi}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial\varphi}{\partial\theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial\varphi}{\partial\phi}\mathbf{e}_\phi \\ \nabla\cdot\mathbf{v} &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 v_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta v_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial\phi} \\ \nabla^2\varphi &= \frac{\partial^2\varphi}{\partial r^2} + \frac{2}{r}\frac{\partial\varphi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial\theta^2} + \frac{\cot\theta}{r^2}\frac{\partial\varphi}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\varphi}{\partial\phi^2}\end{aligned} \tag{1.6.38}$$

1.6.11 The Directional Derivative

Consider a function $\phi(\mathbf{x})$. The directional derivative of φ in the direction of some vector $\mathbf{w}$ is the change in φ in that direction. Now the difference between its values at position $\mathbf{x}$ and $\mathbf{x}+\mathbf{w}$ is, Fig. 1.6.10,

$$d\phi = \phi(\mathbf{x}+\mathbf{w}) - \phi(\mathbf{x}) \tag{1.6.39}$$

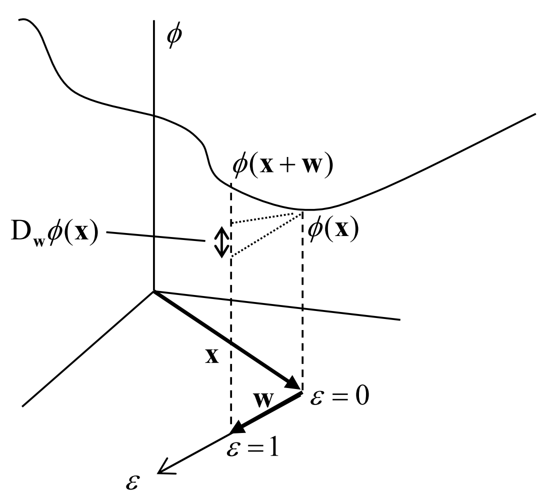


Figure 1.6.10: the directional derivative

An approximation to $d\phi$ can be obtained by introducing a parameter ε and by considering the function $\phi(\mathbf{x}+\varepsilon\mathbf{w})$; one has $\phi(\mathbf{x}+\varepsilon\mathbf{w})_{\varepsilon=0} = \phi(\mathbf{x})$ and $\phi(\mathbf{x}+\varepsilon\mathbf{w})_{\varepsilon=1} = \phi(\mathbf{x}+\mathbf{w})$.

If one treats ϕ as a function of ε, a Taylor's series about $\varepsilon = 0$ gives

$$\phi(\varepsilon) = \phi(0) + \varepsilon \left.\frac{d\phi(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} + \frac{\varepsilon^2}{2}\left.\frac{d^2\phi(\varepsilon)}{d\varepsilon^2}\right|_{\varepsilon=0} + \cdots$$

or, writing it as a function of $\mathbf{x}+\varepsilon\mathbf{w}$,

$$\phi(\mathbf{x}+\varepsilon\mathbf{w}) = \phi(\mathbf{x}) + \varepsilon \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \phi(\mathbf{x}+\varepsilon\mathbf{w}) + \cdots$$

By setting $\varepsilon = 1$, the derivative here can be seen to be a linear approximation to the increment $d\phi$, Eqn. 1.6.39. This is defined as the **directional derivative** of the function $\phi(\mathbf{x})$ at the point **x** in the direction of **w**, and is denoted by

$$\boxed{\partial_{\mathbf{x}}\phi[\mathbf{w}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \phi(\mathbf{x}+\varepsilon\mathbf{w})} \qquad \textbf{The Directional Derivative} \qquad (1.6.40)$$

The directional derivative is also written as $\mathrm{D}_{\mathbf{w}}\phi(\mathbf{x})$.

The power of the directional derivative as defined by Eqn. 1.6.40 is its generality, as seen in the following example.

Example (the Directional Derivative of the Determinant)

Consider the directional derivative of the determinant of the 2×2 matrix **A**, in the direction of a second matrix **T** (the word "direction" is obviously used loosely in this context). One has

$$\begin{aligned}
\partial_{\mathbf{A}}(\det\mathbf{A})[\mathbf{T}] &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \det(\mathbf{A}+\varepsilon\mathbf{T}) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left[(A_{11}+\varepsilon T_{11})(A_{22}+\varepsilon T_{22}) - (A_{12}+\varepsilon T_{12})(A_{21}+\varepsilon T_{21})\right] \\
&= A_{11}T_{22} + A_{22}T_{11} - A_{12}T_{21} - A_{21}T_{12}
\end{aligned}$$

■

The Directional Derivative and The Gradient

Consider a scalar-valued function ϕ of a vector **z**. Let **z** be a function of a parameter ε, $\phi \equiv \phi(z_1(\varepsilon), z_2(\varepsilon), z_3(\varepsilon))$. Then

$$\frac{d\phi}{d\varepsilon} = \frac{\partial\phi}{\partial z_i}\frac{dz_i}{d\varepsilon} = \frac{\partial\phi}{\partial\mathbf{z}}\cdot\frac{d\mathbf{z}}{d\varepsilon}$$

Thus, with $\mathbf{z} = \mathbf{x} + \varepsilon\mathbf{w}$,

$$\partial_{\mathbf{x}}\phi[\mathbf{w}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\phi(\mathbf{z}(\varepsilon)) = \left(\frac{\partial\phi}{\partial\mathbf{z}}\cdot\frac{d\mathbf{z}}{d\varepsilon}\right)_{\varepsilon=0} = \frac{\partial\phi}{\partial\mathbf{x}}\cdot\mathbf{w} \qquad (1.6.41)$$

which can be compared with Eqn. 1.6.11. Note that for Eqns. 1.6.11 and 1.6.41 to be consistent definitions of the directional derivative, **w** here should be a *unit* vector.

1.6.12 Formal Treatment of Vector Calculus

The calculus of vectors is now treated more formally in what follows, following on from the introductory section in §1.2. Consider a vector **h**, an element of the Euclidean vector space *E*, $\mathbf{h} \in E$. In order to be able to speak of limits as elements become "small" or "close" to each other in this space, one requires a norm. Here, take the standard Euclidean norm on *E*, Eqn. 1.2.8,

$$\|\mathbf{h}\| \equiv \sqrt{\langle\mathbf{h},\mathbf{h}\rangle} = \sqrt{\mathbf{h}\cdot\mathbf{h}} \qquad (1.6.42)$$

Consider next a scalar function $f : E \to R$. If there is a constant $M > 0$ such that $|f(\mathbf{h})| \le M\|\mathbf{h}\|$ as $\mathbf{h} \to \mathbf{o}$, then one writes

$$f(\mathbf{h}) = O(\|\mathbf{h}\|) \quad \text{as} \quad \mathbf{h} \to \mathbf{o} \qquad (1.6.43)$$

This is called the **Big Oh** (or **Landau**) notation. Eqn. 1.6.43 states that $|f(\mathbf{h})|$ goes to zero at least as fast as $\|\mathbf{h}\|$. An expression such as

$$f(\mathbf{h}) = g(\mathbf{h}) + O(\|\mathbf{h}\|) \qquad (1.6.44)$$

then means that $|f(\mathbf{h}) - g(\mathbf{h})|$ is smaller than $\|\mathbf{h}\|$ for $\mathbf{h}$ sufficiently close to **o**.

Similarly, if

$$\frac{f(\mathbf{h})}{\|\mathbf{h}\|} \to 0 \quad \text{as} \quad \mathbf{h} \to \mathbf{o} \qquad (1.6.45)$$

then one writes $f(\mathbf{h}) = o(\|\mathbf{h}\|)$ as $\mathbf{h} \to \mathbf{o}$. This implies that $|f(\mathbf{h})|$ goes to zero faster than $\|\mathbf{h}\|$.

A **field** is a function which is defined in a Euclidean (point) space E^3. A **scalar field** is then a function $f : E^3 \to R$. A scalar field is **differentiable** at a point $\mathbf{x} \in E^3$ if there exists a vector $Df(\mathbf{x}) \in E$ such that

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + Df(\mathbf{x})\cdot\mathbf{h} + o(\|\mathbf{h}\|) \quad \text{for all} \quad \mathbf{h} \in E \tag{1.6.46}$$

In that case, the vector $Df(\mathbf{x})$ is called the **derivative** (or **gradient**) of f at $\mathbf{x}$ (and is given the symbol $\nabla f(\mathbf{x})$).

Now setting $\mathbf{h} = \varepsilon\mathbf{w}$ in 1.6.46, where $\mathbf{w} \in E$ is a unit vector, dividing through by ε and taking the limit as $\varepsilon \to 0$, one has the equivalent statement

$$\nabla f(\mathbf{x})\cdot\mathbf{w} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} f(\mathbf{x}+\varepsilon\mathbf{w}) \quad \text{for all} \quad \mathbf{w} \in E \tag{1.6.47}$$

which is 1.6.41. In other words, for the derivative to exist, the scalar field must have a directional derivative in all directions at $\mathbf{x}$.

Using the chain rule as in §1.6.11, Eqn. 1.6.47 can be expressed in terms of the Cartesian basis $\{\mathbf{e}_i\}$,

$$\nabla f(\mathbf{x})\cdot\mathbf{w} = \frac{\partial f}{\partial x_i} w_i = \frac{\partial f}{\partial x_i}\mathbf{e}_i \cdot w_j\mathbf{e}_j \tag{1.6.48}$$

This must be true for all $\mathbf{w}$ and so, in a Cartesian basis,

$$\nabla f(\mathbf{x}) = \frac{\partial f}{\partial x_i}\mathbf{e}_i \tag{1.6.49}$$

which is Eqn. 1.6.9.

1.6.13 Problems

1. A particle moves along a curve in space defined by
$$\mathbf{r} = (t^3 - 4t)\mathbf{e}_1 + (t^2 + 4t)\mathbf{e}_2 + (8t^2 - 3t^3)\mathbf{e}_3$$
Here, t is time. Find
(i) a unit tangent vector at $t = 2$
(ii) the magnitudes of the tangential and normal components of acceleration at $t = 2$
2. Use the index notation (1.3.12) to show that $\frac{d}{dt}(\mathbf{v}\times\mathbf{a}) = \mathbf{v}\times\frac{d\mathbf{a}}{dt} + \frac{d\mathbf{v}}{dt}\times\mathbf{a}$. Verify this result for $\mathbf{v} = 3t\mathbf{e}_1 - t^2\mathbf{e}_3$, $\mathbf{a} = t^2\mathbf{e}_1 + t\mathbf{e}_2$. [Note: the permutation symbol and the unit vectors are independent of t; the components of the vectors are scalar functions of t which can be differentiated in the usual way, for example by using the product rule of differentiation.]

3. The density distribution throughout a material is given by $\rho = 1 + \mathbf{x} \cdot \mathbf{x}$.
   (i) what sort of function is this?
   (ii) the density is given in symbolic notation - write it in index notation
   (iii) evaluate the gradient of ρ
   (iv) give a unit vector in the direction in which the density is increasing the most
   (v) give a unit vector in *any* direction in which the density is not increasing
   (vi) take any unit vector other than the base vectors and the other vectors you used above and calculate $d\rho / dx$ in the direction of this unit vector
   (vii) evaluate and sketch all these quantities for the point (2,1).
   In parts (iii-iv), give your answer in (a) symbolic, (b) index, and (c) full notation.
4. Consider the scalar field defined by $\phi = x^2 + 3yx + 2z$.
   (i) find the unit normal to the surface of constant ϕ at the origin (0,0,0)
   (ii) what is the maximum value of the directional derivative of ϕ at the origin?
   (iii) evaluate $d\phi / dx$ at the origin if $d\mathbf{x} = ds(\mathbf{e}_1 + \mathbf{e}_3)$.
5. If $\mathbf{u} = x_1 x_2 x_3 \mathbf{e}_1 + x_1 x_2 \mathbf{e}_2 + x_1 \mathbf{e}_3$, determine $\operatorname{div} \mathbf{u}$ and $\operatorname{curl} \mathbf{u}$.
6. Determine the constant *a* so that the vector
$$\mathbf{v} = (x_1 + 3x_2)\mathbf{e}_1 + (x_2 - 2x_3)\mathbf{e}_2 + (x_1 + ax_3)\mathbf{e}_3$$
   is solenoidal.
7. Show that $\operatorname{curl} \mathbf{v} = 2\boldsymbol{\omega}$ (see also Problem 9 in §1.1).
8. Verify the identities (1.6.25-26).
9. Use (1.6.14) to derive the Nabla operator in cylindrical coordinates (1.6.30).
10. Derive Eqn. (1.6.34), the curl of a vector and the Laplacian of a scalar in the cylindrical coordinates.
11. Derive (1.6.38), the gradient, divergence and Laplacian in spherical coordinates.
12. Show that the directional derivative $\mathrm{D}_{\mathbf{v}} \phi(\mathbf{u})$ of the scalar-valued function of a vector $\phi(\mathbf{u}) = \mathbf{u} \cdot \mathbf{u}$, in the direction **v**, is $2\mathbf{u} \cdot \mathbf{v}$.
13. Show that the directional derivative of the functional
$$U(v(x)) = \frac{1}{2} \int_0^l EI \left(\frac{d^2 v}{dx^2} \right)^2 dx - \int_0^l p(x) v(x) dx$$
   in the direction of $\omega(x)$ is given by
$$\int_0^l EI \frac{d^2 v(x)}{dx^2} \frac{d^2 \omega(x)}{dx^2} dx - \int_0^l p(x) \omega(x) dx .$$

1.7 Vector Calculus 2 - Integration

1.7.1 Ordinary Integrals of a Vector

A vector can be integrated in the ordinary way to produce another vector, for example

$$\int_1^2 \left\{ \left(t - t^2\right)\mathbf{e}_1 + 2t^2\mathbf{e}_2 - 3\mathbf{e}_3 \right\} dt = -\frac{5}{6}\mathbf{e}_1 + \frac{15}{2}\mathbf{e}_2 - 3\mathbf{e}_3$$

1.7.2 Line Integrals

Discussed here is the notion of a definite integral involving a vector function that generates a scalar.

Let $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ be a position vector tracing out the curve C between the points p_1 and p_2. Let $\mathbf{f}$ be a vector field. Then

$$\int_{p_1}^{p_2} \mathbf{f} \cdot d\mathbf{x} = \int_C \mathbf{f} \cdot d\mathbf{x} = \int_C \left\{ f_1 dx_1 + f_2 dx_2 + f_3 dx_3 \right\}$$

is an example of a line integral.

Example (of a Line Integral)

A particle moves along a path C from the point $(0,0,0)$ to $(1,1,1)$, where C is the straight line joining the points, Fig. 1.7.1. The particle moves in a force field given by

$$\mathbf{f} = \left(3x_1^2 + 6x_2\right)\mathbf{e}_1 - 14x_2x_3\mathbf{e}_2 + 20x_1x_3^2\mathbf{e}_3$$

What is the work done on the particle?

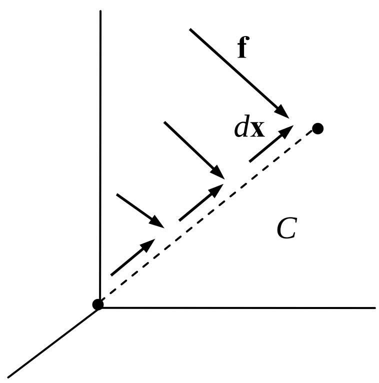


Figure 1.7.1: a particle moving in a force field

Solution
The work done is

$$W = \int_C \mathbf{f} \cdot d\mathbf{x} = \int_C \left\{ \left(3x_1^2 + 6x_2\right) dx_1 - 14x_2x_3 dx_2 + 20x_1x_3^2 dx_3 \right\}$$

The straight line can be written in the parametric form $x_1 = t, x_2 = t, x_3 = t$, so that

$$W = \int_0^1 \left(20t^3 - 11t^2 + 6t\right) dt = \frac{13}{3} \quad \text{or} \quad W = \int_C \mathbf{f} \cdot \frac{d\mathbf{x}}{dt} dt = \int_C \mathbf{f} \cdot \left(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3\right) dt = \frac{13}{3}$$

■

If C is a closed curve, i.e. a loop, the line integral is often denoted $\oint_C \mathbf{v} \cdot d\mathbf{x}$.

Note: in fluid mechanics and aerodynamics, when **v** is the velocity field, this integral $\oint_C \mathbf{v} \cdot d\mathbf{x}$ is called the **circulation** of **v** about C.

1.7.3 Conservative Fields

If for a vector **f** one can find a scalar ϕ such that

$$\mathbf{f} = \nabla \phi \tag{1.7.1}$$

then

(1) $\int_{p_1}^{p_2} \mathbf{f} \cdot d\mathbf{x}$ is independent of the path C joining p_1 and p_2

(2) $\oint_C \mathbf{f} \cdot d\mathbf{x} = 0$ around any closed curve C

In such a case, **f** is called a **conservative vector field** and ϕ is its **scalar potential**[1]. For example, the work done by a conservative force field **f** is

$$\int_{p_1}^{p_2} \mathbf{f} \cdot d\mathbf{x} = \int_{p_1}^{p_2} \nabla \phi \cdot d\mathbf{x} = \int_{p_1}^{p_2} \frac{\partial \phi}{\partial x_i} dx_i = \int_{p_1}^{p_2} d\phi = \phi(p_2) - \phi(p_1)$$

which clearly depends only on the values at the end-points p_1 and p_2, and not on the path taken between them.

It can be shown that a vector **f** is conservative if and only if $\text{curl}\mathbf{f} = \mathbf{o}$ {▲Problem 3}.

[1] in general, of course, there does not exist a scalar field ϕ such that $\mathbf{f} = \nabla \phi$; this is not surprising since a vector field has three scalar components whereas $\nabla \phi$ is determined from just one

Example (of a Conservative Force Field)

The gravitational force field $\mathbf{f} = -mg\mathbf{e}_3$ is an example of a conservative vector field. Clearly, $\operatorname{curl}\mathbf{f} = \mathbf{o}$, and the gravitational scalar potential is $\phi = -mgx_3$:

$$W = -\int_{p_1}^{p_2} mg\mathbf{e}_3 \cdot d\mathbf{x} = -mg\int_{p_1}^{p_2} dx_3 = -mg[x_3(p_2) - x_3(p_1)] = \phi(p_2) - \phi(p_1)$$

■

Example (of a Conservative Force Field)

Consider the force field

$$\mathbf{f} = (2x_1x_2 + x_3^3)\mathbf{e}_1 + x_1^2\mathbf{e}_2 + 3x_1x_3^2\mathbf{e}_3$$

Show that it is a conservative force field, find its scalar potential and find the work done in moving a particle in this field from $(1,-2,1)$ to $(3,1,4)$.

<u>Solution</u>

One has

$$\operatorname{curl}\mathbf{f} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial/\partial x_1 & \partial/\partial x_2 & \partial/\partial x_3 \\ 2x_1x_2 + x_3^3 & x_1^2 & 3x_1x_3^2 \end{vmatrix} = \mathbf{o}$$

so the field is conservative.

To determine the scalar potential, let

$$f_1\mathbf{e}_1 + f_2\mathbf{e}_2 + f_3\mathbf{e}_3 = \frac{\partial\phi}{\partial x_1}\mathbf{e}_1 + \frac{\partial\phi}{\partial x_2}\mathbf{e}_2 + \frac{\partial\phi}{\partial x_3}\mathbf{e}_3.$$

Equating coefficients and integrating leads to

$$\begin{aligned}
\phi &= x_1^2x_2 + x_1x_3^3 + p(x_2, x_3) \\
\phi &= x_1^2x_2 \qquad\quad + q(x_1, x_3) \\
\phi &= \qquad\quad x_1x_3^3 + r(x_1, x_1)
\end{aligned}$$

which agree if one chooses $p = 0,\ q = x_1x_3^3,\ r = x_1^2x_2$, so that $\phi = x_1^2x_2 + x_1x_3^3$, to which may be added a constant.

The work done is

$$W = \phi(3,1,4) - \phi(1,-2,1) = 202$$

■

Helmholtz Theory

As mentioned, a conservative vector field which is irrotational, i.e. $\mathbf{f} = \nabla\phi$, implies $\nabla \times \mathbf{f} = \mathbf{o}$, and *vice versa*. Similarly, it can be shown that if one can find a vector **a** such that $\mathbf{f} = \nabla \times \mathbf{a}$, where **a** is called the **vector potential**, then **f** is solenoidal, i.e. $\nabla \cdot \mathbf{f} = 0$ {▲Problem 4}.

Helmholtz showed that a vector can always be represented in terms of a scalar potential ϕ and a vector potential **a**:[2]

Type of Vector	Condition	Representation
General		$\mathbf{f} = \nabla\phi + \nabla \times \mathbf{a}$
Irrotational (conservative)	$\nabla \times \mathbf{f} = \mathbf{o}$	$\mathbf{f} = \nabla\phi$
Solenoidal	$\nabla \cdot \mathbf{f} = 0$	$\mathbf{f} = \nabla \times \mathbf{a}$

1.7.4 Double Integrals

The most elementary type of two-dimensional integral is that over a plane region. For example, consider the integral over a region *R* in the $x_1 - x_2$ plane, Fig. 1.7.2. The integral

$$\iint_R dx_1 dx_2$$

then gives the area of *R* and, just as the one dimensional integral of a function gives the area under the curve, the integral

$$\iint_R f(x_1, x_2) dx_1 dx_2$$

gives the volume under the (in general, curved) surface $x_3 = f(x_1, x_2)$. These integrals are called **double integrals**.

[2] this decomposition can be made unique by requiring that $\mathbf{f} \to \mathbf{0}$ as $\mathbf{x} \to \infty$; in general, if one is given **f**, then ϕ and **a** can be obtained by solving a number of differential equations

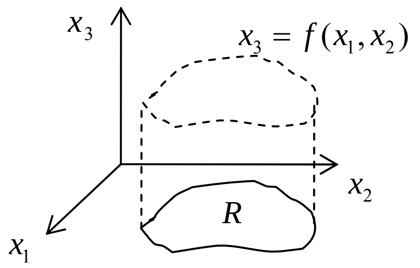


Figure 1.7.2: integration over a region

Change of variables in Double Integrals

To evaluate integrals of the type $\iint_R f(x_1, x_2) dx_1 dx_2$, it is often convenient to make a change of variable. To do this, one must find an elemental surface area in terms of the new variables, t_1, t_2 say, equivalent to that in the x_1, x_2 coordinate system, $dS = dx_1 dx_2$.

The region *R* over which the integration takes place is the plane surface $g(x_1, x_2) = 0$. Just as a curve can be represented by a position vector of one single parameter *t* (*cf.* §1.6.2), this surface can be represented by a position vector with two parameters[3], t_1 and t_2:

$$\mathbf{x} = x_1(t_1, t_2)\mathbf{e}_1 + x_2(t_1, t_2)\mathbf{e}_2$$

Parameterising the plane surface in this way, one can calculate the element of surface dS in terms of t_1, t_2 by considering curves of constant t_1, t_2, as shown in Fig. 1.7.3. The vectors bounding the element are

$$d\mathbf{x}^{(1)} = d\mathbf{x}\big|_{t_2 \text{ const}} = \frac{\partial \mathbf{x}}{\partial t_1} dt_1, \qquad d\mathbf{x}^{(2)} = d\mathbf{x}\big|_{t_1 \text{ const}} = \frac{\partial \mathbf{x}}{\partial t_2} dt_2 \tag{1.7.2}$$

so the area of the element is given by

$$dS = \left| d\mathbf{x}^{(1)} \times d\mathbf{x}^{(2)} \right| = \left| \frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right| dt_1 dt_2 = J\, dt_1 dt \tag{1.7.3}$$

where *J* is the **Jacobian** of the transformation,

[3] for example, the unit circle $x_1^2 + x_2^2 - 1 = 0$ can be represented by $\mathbf{x} = t_1 \cos t_2 \mathbf{e}_1 + t_1 \sin t_2 \mathbf{e}_2$, $0 < t_1 \le 1$, $0 < t_2 \le 2\pi$ (t_1, t_2 being in this case the polar coordinates r, θ, respectively)

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_2}{\partial t_1} \\ \dfrac{\partial x_1}{\partial t_2} & \dfrac{\partial x_2}{\partial t_2} \end{vmatrix} \quad \text{or} \quad J = \begin{vmatrix} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_1}{\partial t_2} \\ \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_2}{\partial t_2} \end{vmatrix} \tag{1.7.4}$$

The Jacobian is also often written using the notation

$$dx_1 dx_2 = J dt_1 dt_2, \qquad J = \left| \frac{\partial(x_1, x_2)}{\partial(t_1, t_2)} \right|$$

The integral can now be written as

$$\iint_R f(t_1, t_2) J dt_1 dt_2$$

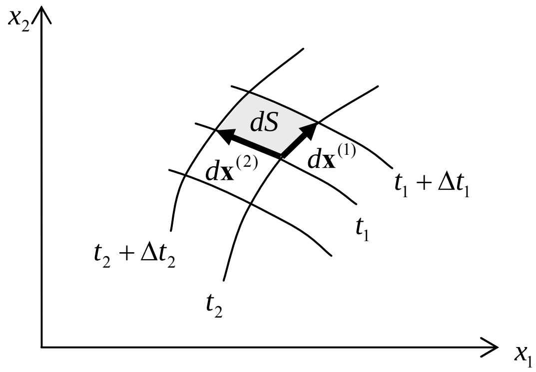


Figure 1.7.3: a surface element

Example

Consider a region R, the quarter unit-circle in the first quadrant, $0 \le x_2 \le \sqrt{1 - x_1^2}$, $0 \le x_1 \le 1$. The moment of inertia about the x_1 – axis is defined by

$$I_{x_1} \equiv \iint_R x_2^2 dx_1 dx_2$$

Transform the integral into the new coordinate system t_1, t_2 by making the substitutions[4] $x_1 = t_1 \cos t_2$, $x_2 = t_1 \sin t_2$. Then

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_1}{\partial t_2} \\ \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_2}{\partial t_2} \end{vmatrix} = \begin{vmatrix} \cos t_2 & -t_1 \sin t_2 \\ \sin t_2 & t_1 \cos t_2 \end{vmatrix} = t_1$$

[4] these are the polar coordinates, t_1, t_2 equal to r, θ, respectively

so

$$I_{x_1} = \int_0^{\pi/2}\int_0^1 t_1^3 \sin^2 t_2 dt_1 dt_2 = \frac{\pi}{16}$$

■

1.7.5 Surface Integrals

Up to now, double integrals over a plane region have been considered. In what follows, consideration is given to integrals over more complex, curved, surfaces in space, such as the surface of a sphere.

Surfaces

Again, a curved surface can be parameterized by t_1, t_2, now by the position vector

$$\mathbf{x} = x_1(t_1,t_2)\mathbf{e}_1 + x_2(t_1,t_2)\mathbf{e}_2 + x_3(t_1,t_2)\mathbf{e}_3$$

One can generate a curve *C* on the surface *S* by taking $t_1 = t_1(s)$, $t_2 = t_2(s)$ so that *C* has position vector, Fig. 1.7.4,

$$\mathbf{x}(s) = \mathbf{x}(t_1(s), t_2(s))$$

A vector tangent to *C* at a point *p* on *S* is, from Eqn. 1.6.3,

$$\frac{d\mathbf{x}}{ds} = \frac{\partial \mathbf{x}}{\partial t_1}\frac{dt_1}{ds} + \frac{\partial \mathbf{x}}{\partial t_2}\frac{dt_2}{ds}$$

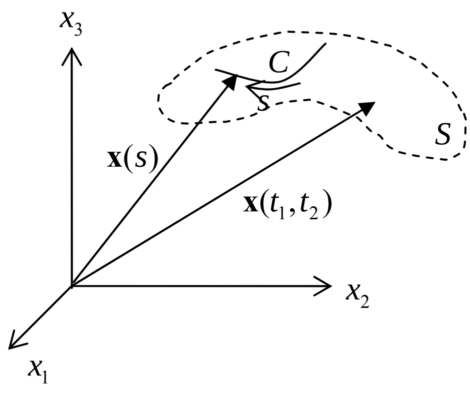


Figure 1.7.4: a curved surface

Many different curves *C* pass through *p*, and hence there are many different tangents, with different corresponding values of $dt_1/ds, dt_2/ds$. Thus the partial derivatives $\partial\mathbf{x}/\partial t_1, \partial\mathbf{x}/\partial t_2$ must also both be tangential to *C* and so a normal to the surface at *p* is given by their cross-product, and a unit normal is

$$\mathbf{n} = \left(\frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right) / \left| \frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right| \tag{1.7.5}$$

In some cases, it is possible to use a non-parametric form for the surface, for example $g(x_1, x_2, x_3) = c$, in which case the normal can be obtained simply from $\mathbf{n} = \mathrm{grad}\, g / \left| \mathrm{grad}\, g \right|$.

Example (Parametric Representation and the Normal to a Sphere)

The surface of a sphere of radius a can be parameterised as[5]

$$\mathbf{x} = a\left\{ \sin t_1 \cos t_2 \mathbf{e}_1 + \sin t_1 \sin t_2 \mathbf{e}_2 + \cos t_1 \mathbf{e}_3 \right\}, \qquad 0 \le t_1 \le \pi,\ 0 \le t_2 \le 2\pi$$

Here, lines of $t_1 = \text{const}$ are parallel to the $x_1 - x_2$ plane ("parallels"), whereas lines of $t_2 = \text{const}$ are "meridian" lines, Fig. 1.7.5. If one takes the simple expressions $t_1 = s, t_2 = \pi/2 - s$, over $0 \le s \le \pi/2$, one obtains a curve C_1 joining $(0,0,1)$ and $(1,0,0)$, and passing through $(1/2, 1/2, 1/\sqrt{2})$, as shown.

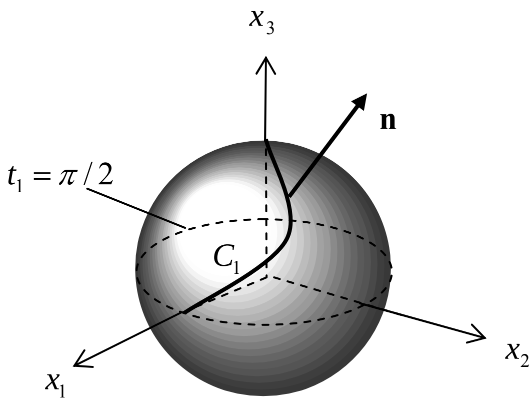


Figure 1.7.5: a sphere

The partial derivatives with respect to the parameters are

$$\frac{\partial \mathbf{x}}{\partial t_1} = a\left\{ \cos t_1 \cos t_2 \mathbf{e}_1 + \cos t_1 \sin t_2 \mathbf{e}_2 - \sin t_1 \mathbf{e}_3 \right\}$$
$$\frac{\partial \mathbf{x}}{\partial t_2} = a\left\{ -\sin t_1 \sin t_2 \mathbf{e}_1 + \sin t_1 \cos t_2 \mathbf{e}_2 \right\}$$

so that

$$\frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} = a^2 \left\{ \sin^2 t_1 \cos t_2 \mathbf{e}_1 + \sin^2 t_1 \sin t_2 \mathbf{e}_2 + \sin t_1 \cos t_1 \mathbf{e}_3 \right\}$$

[5] these are the **spherical coordinates** (see §1.6.10); $t_1 = \theta, t_2 = \phi$

and a unit normal to the spherical surface is

$$\mathbf{n} = \sin t_1 \cos t_2 \mathbf{e}_1 + \sin t_1 \sin t_2 \mathbf{e}_2 + \cos t_1 \mathbf{e}_3$$

For example, at $t_1 = t_2 = \pi/4$ (this is on the curve C_1), one has

$$\mathbf{n}(\pi/4, \pi/4) = \tfrac{1}{2}\mathbf{e}_1 + \tfrac{1}{2}\mathbf{e}_2 + \tfrac{1}{\sqrt{2}}\mathbf{e}_3$$

and, as expected, it is in the same direction as **r**.

■

Surface Integrals

Consider now the integral $\iint_S \mathbf{f} dS$ where **f** is a vector function and *S* is some curved surface. As for the integral over the plane region,

$$dS = \left| d\mathbf{x}\big|_{t_2 \text{ const}} \times d\mathbf{x}\big|_{t_1 \text{ const}} \right| = \left| \frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right| dt_1 dt_2 ,$$

only now dS is not "flat" and **x** is three dimensional. The integral can be evaluated if one parameterises the surface with t_1, t_2 and then writes

$$\iint_S \mathbf{f} \left| \frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right| dt_1 dt_2$$

One way to evaluate this cross product is to use the relation (**Lagrange's identity**, Problem 15, §1.3)

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \tag{1.7.6}$$

so that

$$\left| \frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right|^2 = \left(\frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right) \cdot \left(\frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right) = \left(\frac{\partial \mathbf{x}}{\partial t_1} \cdot \frac{\partial \mathbf{x}}{\partial t_1} \right)\left(\frac{\partial \mathbf{x}}{\partial t_2} \cdot \frac{\partial \mathbf{x}}{\partial t_2} \right) - \left(\frac{\partial \mathbf{x}}{\partial t_1} \cdot \frac{\partial \mathbf{x}}{\partial t_2} \right)^2 \tag{1.7.7}$$

Example (Surface Area of a Sphere)

Using the parametric form for a sphere given above, one obtains

$$\left| \frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right|^2 = a^4 \sin^2 t_1$$

so that

$$\text{area} = \iint_S dS = a^2 \int_0^{2\pi}\int_0^{\pi} \sin t_1 dt_1 dt_2 = 4\pi a^2$$

■

Flux Integrals

Surface integrals often involve the normal to the surface, as in the following example.

Example

If $\mathbf{f} = 4x_1x_3\mathbf{e}_1 - x_2^2\mathbf{e}_2 + x_2x_3\mathbf{e}_3$, evaluate $\iint_S \mathbf{f}\cdot\mathbf{n}dS$, where S is the surface of the cube bounded by $x_1 = 0,1;\ x_2 = 0,1;\ x_3 = 0,1$, and $\mathbf{n}$ is the unit outward normal, Fig. 1.7.6.

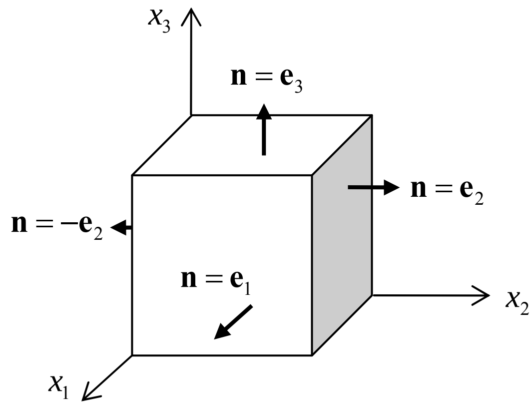


Figure 1.7.6: the unit cube

<u>Solution</u>

The integral needs to be evaluated over the six faces. For the face with $\mathbf{n} = +\mathbf{e}_1$, $x_1 = 1$ and

$$\iint_S \mathbf{f}\cdot\mathbf{n}dS = \int_0^1\int_0^1 \left(4x_3\mathbf{e}_1 - x_2^2\mathbf{e}_2 + x_2x_3\mathbf{e}_3\right)\cdot\mathbf{e}_1 dx_2dx_3 = 4\int_0^1\int_0^1 x_3dx_2dx_3 = 2$$

Similarly for the other five sides, whence $\iint_S \mathbf{f}\cdot\mathbf{n}dS = \frac{3}{2}$.

■

Integrals of the form $\iint_S \mathbf{f}\cdot\mathbf{n}dS$ are known as **flux integrals** and arise quite often in applications. For example, consider a material flowing with velocity **v**, in particular the flow through a small surface element dS with outward unit normal **n**, Fig. 1.7.7. The volume of material flowing through the surface in time dt is equal to the volume of the slanted cylinder shown, which is the base dS times the height. The slanted height is (=

velocity × time) is $|\mathbf{v}|dt$, and the vertical height is then $\mathbf{v}\cdot\mathbf{n}dt$. Thus the *rate* of flow is the **volume flux** (volume per unit time) through the surface element: $\mathbf{v}\cdot\mathbf{n}dS$.

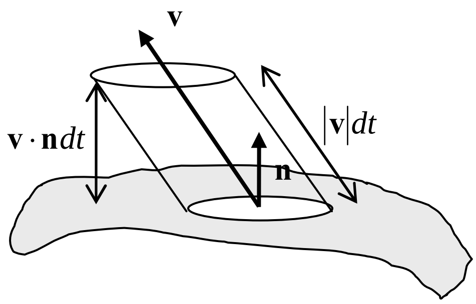


Figure 1.7.7: flow through a surface element

The total (volume) flux *out* of a surface S is then[6]

$$\textit{volume flux:} \quad \iint_S \mathbf{v}\cdot\mathbf{n}dS \tag{1.7.8}$$

Similarly, the **mass flux** is given by

$$\textit{mass flux:} \quad \iint_S \rho\mathbf{v}\cdot\mathbf{n}dS \tag{1.7.9}$$

For more complex surfaces, one can write using Eqn. 1.7.3, 1.7.5,

$$\iint_S \mathbf{f}\cdot\mathbf{n}dS = \iint_S \mathbf{f}\cdot\left(\frac{\partial\mathbf{x}}{\partial t_1}\times\frac{\partial\mathbf{x}}{\partial t_2}\right)dt_1 dt_2$$

Example (of a Flux Integral)

Compute the flux integral $\iint_S \mathbf{f}\cdot\mathbf{n}dS$, where S is the parabolic cylinder represented by

$$x_2 = x_1^2, \quad 0 \le x_1 \le 2, \quad 0 \le x_3 \le 3$$

and $\mathbf{f} = x_2\mathbf{e}_1 + 2\mathbf{e}_2 + x_1 x_3\mathbf{e}_3$, Fig. 1.7.8.

<u>Solution</u>

Making the substitutions $x_1 = t_1$, $x_3 = t_2$, so that $x_2 = t_1^2$, the surface can be represented by the position vector

[6] if **v** acts in the same direction as **n**, i.e. pointing outward, the dot product is positive and this integral is positive; if, on the other hand, material is flowing *in* through the surface, **v** and **n** are in opposite directions and the dot product is negative, so the integral is negative

$$\mathbf{x} = t_1\mathbf{e}_1 + t_1^2\mathbf{e}_2 + t_2\mathbf{e}_3, \qquad 0 \le t_1 \le 2,\ 0 \le t_2 \le 3$$

Then $\partial\mathbf{x}/\partial t_1 = \mathbf{e}_1 + 2t_1\mathbf{e}_2$, $\partial\mathbf{x}/\partial t_2 = \mathbf{e}_3$ and

$$\frac{\partial\mathbf{x}}{\partial t_1} \times \frac{\partial\mathbf{x}}{\partial t_2} = 2t_1\mathbf{e}_1 - \mathbf{e}_2$$

so the integral becomes

$$\int_0^3\int_0^2 \left(t_1^2\mathbf{e}_1 + 2\mathbf{e}_2 + t_1t_2\mathbf{e}_3\right)\cdot\left(2t_1\mathbf{e}_1 - \mathbf{e}_2\right)dt_1dt_2 = 12$$

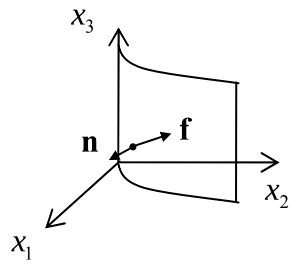


Figure 1.7.8: flux through a parabolic cylinder

Note: in this example, the value of the integral depends on the choice of **n**. If one chooses $-\mathbf{n}$ instead of **n**, one would obtain -12. The normal in the opposite direction (on the "other side" of the surface) can be obtained by simply switching t_1 and t_2, since $\partial\mathbf{x}/\partial t_1 \times \partial\mathbf{x}/\partial t_2 = -\partial\mathbf{x}/\partial t_2 \times \partial\mathbf{x}/\partial t_1$.

■

Surface flux integrals can also be evaluated by first converting them into double integrals over a plane region. For example, if a surface *S* has a projection *R* on the $x_1 - x_2$ plane, then an element of surface dS is related to the projected element dx_1dx_2 through (see Fig. 1.7.9)

$$\cos\theta dS = (\mathbf{n}\cdot\mathbf{e}_3)dS = dx_1dx_2$$

and so

$$\iint_S \mathbf{f}\cdot\mathbf{n}dS = \iint_R \mathbf{f}\cdot\mathbf{n}\frac{1}{|\mathbf{n}\cdot\mathbf{e}_3|}dx_1dx_2$$

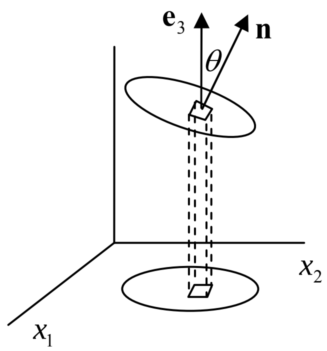


Figure 1.7.9: projection of a surface element onto a plane region

The Normal and Surface Area Elements

It is sometimes convenient to associate a special vector $d\mathbf{S}$ with a differential element of surface area dS, where

$$d\mathbf{S} = \mathbf{n}\, dS$$

so that $d\mathbf{S}$ is the vector with magnitude dS and direction of the unit normal to the surface. Flux integrals can then be written as

$$\iint_S \mathbf{f} \cdot \mathbf{n} dS = \iint_S \mathbf{f} \cdot d\mathbf{S}$$

1.7.6 Volume Integrals

The volume integral, or triple integral, is a generalisation of the double integral.

Change of Variable in Volume Integrals

For a volume integral, it is often convenient to make the change of variables $(x_1, x_2, x_3) \to (t_1, t_2, t_3)$. The volume of an element dV is given by the triple scalar product (Eqns. 1.1.5, 1.3.17)

$$dV = \left(\frac{\partial \mathbf{x}}{\partial t_1} \times \frac{\partial \mathbf{x}}{\partial t_2} \right) \cdot \frac{\partial \mathbf{x}}{\partial t_3} dt_1 dt_2 dt_3 = J dt_1 dt_2 dt_3 \tag{1.7.10}$$

where the Jacobian is now

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_3}{\partial t_1} \\ \dfrac{\partial x_1}{\partial t_2} & \dfrac{\partial x_2}{\partial t_2} & \dfrac{\partial x_3}{\partial t_2} \\ \dfrac{\partial x_1}{\partial t_3} & \dfrac{\partial x_2}{\partial t_3} & \dfrac{\partial x_3}{\partial t_3} \end{vmatrix} \quad \text{or} \quad J = \begin{vmatrix} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_1}{\partial t_2} & \dfrac{\partial x_1}{\partial t_3} \\ \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_2}{\partial t_2} & \dfrac{\partial x_2}{\partial t_3} \\ \dfrac{\partial x_3}{\partial t_1} & \dfrac{\partial x_3}{\partial t_2} & \dfrac{\partial x_3}{\partial t_3} \end{vmatrix} \tag{1.7.11}$$

so that

$$\iiint_V \mathbf{f}(x_1, x_2, x_3)dxdydz = \iiint_V \mathbf{f}(x_1(t_1, t_2, t_3), x_2(t_1, t_2, t_3), x_3(t_1, t_2, t_3))J\, dt_1 dt_2 dt_3$$

1.7.7 Integral Theorems

A number of integral theorems and relations are presented here (without proof), the most important of which is the divergence theorem. These theorems can be used to simplify the evaluation of line, double, surface and triple integrals. They can also be used in various proofs of other important results.

The Divergence Theorem

Consider an arbitrary differentiable vector field $\mathbf{v}(\mathbf{x}, t)$ defined in some finite region of physical space. Let V be a volume in this space with a closed surface S bounding the volume, and let the outward normal to this bounding surface be $\mathbf{n}$. The **divergence theorem of Gauss** states that (in symbolic and index notation)

$$\boxed{\int_S \mathbf{v} \cdot \mathbf{n} dS = \int_V \operatorname{div} \mathbf{v}\, dV \qquad \int_S v_i n_i dS = \int_V \frac{\partial v_i}{\partial x_i} dV} \quad \textbf{Divergence Theorem} \tag{1.7.12}$$

and one has the following useful identities {▲Problem 10}

$$\begin{aligned} \int_S \phi \mathbf{u} \cdot \mathbf{n} dS &= \int_V \operatorname{div}(\phi \mathbf{u})\, dV \\ \int_S \phi \mathbf{n} dS &= \int_V \operatorname{grad} \phi\, dV \\ \int_S \mathbf{n} \times \mathbf{u} dS &= \int_V \operatorname{curl} \mathbf{u}\, dV \end{aligned} \tag{1.7.13}$$

By applying the divergence theorem to a very small volume, one finds that

$$\operatorname{div} \mathbf{v} = \lim_{V \to 0} \frac{\int_S \mathbf{v} \cdot \mathbf{n} dS}{V}$$

that is, the divergence is equal to the outward flux per unit volume, the result 1.6.18.

Stoke's Theorem

Stoke's theorem transforms line integrals into surface integrals and *vice versa*. It states that

$$\iint_S (\mathrm{curl}\mathbf{f})\cdot \mathbf{n}dS = \oint_C \mathbf{f}\cdot \boldsymbol{\tau}ds \tag{1.7.14}$$

Here C is the boundary of the surface S, $\mathbf{n}$ is the unit outward normal and $\boldsymbol{\tau} = d\mathbf{r}/ds$ is the unit tangent vector.

As has been seen, Eqn. 1.6.24, the curl of the velocity field is a measure of how much a fluid is rotating. The direction of this vector is along the direction of the local axis of rotation and its magnitude measures the local angular velocity of the fluid. Stoke's theorem then states that the amount of rotation of a fluid can be measured by integrating the tangential velocity around a curve (the line integral), or by integrating the amount of vorticity "moving through" a surface bounded by the same curve.

Green's Theorem and Related Identities

Green's theorem relates a line integral to a double integral, and states that

$$\oint_C \{\psi_1 dx_1 + \psi_2 dx_2\} = \iint_R \left(\frac{\partial \psi_2}{\partial x_1} - \frac{\partial \psi_1}{\partial x_2}\right) dx_1 dx_2 \,, \tag{1.7.15}$$

where R is a region in the $x_1 - x_2$ plane bounded by the curve C. In vector form, Green's theorem reads as

$$\oint_C \mathbf{f}\cdot d\mathbf{x} = \iint_R \mathrm{curl}\mathbf{f}\cdot \mathbf{e}_3 dx_1 dx_2 \quad \text{where} \quad \mathbf{f} = \psi_1\mathbf{e}_1 + \psi_2\mathbf{e}_2 \tag{1.7.16}$$

from which it can be seen that Green's theorem is a special case of Stoke's theorem, for the case of a plane surface (region) in the $x_1 - x_2$ plane.

It can also be shown that (this is **Green's first identity**)

$$\iint_S \psi(\mathbf{n}\cdot \mathrm{grad}\phi)dS = \iiint_V \{\psi\nabla^2\phi + \mathrm{grad}\psi\cdot\mathrm{grad}\phi\}dV \tag{1.7.17}$$

Note that the term $\mathbf{n}\cdot\mathrm{grad}\phi$ is the directional derivative of ϕ in the direction of the outward unit normal. This is often denoted as $\partial\phi/\partial n$. Green's first identity can be regarded as a multi-dimensional "integration by parts" – compare the rule $\int u dv = uv - \int v du$ with the identity re-written as

$$\iiint_V \psi(\nabla\cdot\nabla\phi)dV = \iint_S \psi(\nabla\phi\cdot\mathbf{n})dS - \iiint_V (\nabla\psi)\cdot(\nabla\phi)dV \tag{1.7.18}$$

or

$$\iiint_V \psi(\nabla\cdot\mathbf{u})dV = \iint_S \psi(\mathbf{u}\cdot\mathbf{n})dS - \iiint_V (\nabla\psi)\cdot\mathbf{u}dV \tag{1.7.18}$$

One also has the relation (this is **Green's second identity**)

$$\iint_S \{\psi(\mathbf{n}\cdot\text{grad}\phi)-\phi(\mathbf{n}\cdot\text{grad}\psi)\}dS = \iiint_V \{\psi\nabla^2\phi-\phi\nabla^2\psi\}dV \tag{1.7.19}$$

1.7.8 Problems

1. Find the work done in moving a particle in a force field given by $\mathbf{f} = 3x_1x_2\mathbf{e}_1 - 5x_3\mathbf{e}_2 + 10x_1\mathbf{e}_1$ along the curve $x_1 = t^2+1$, $x_2 = 2t^2$, $x_3 = t^3$, from $t=1$ to $t=2$. (Plot the curve.)
2. Show that the following vectors are conservative and find their scalar potentials:
   (i) $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$
   (ii) $\mathbf{v} = e^{x_1x_2}(x_2\mathbf{e}_1 + x_1\mathbf{e}_2)$
   (iii) $\mathbf{u} = (1/x_2)\mathbf{e}_1 - (x_1/x_2^2)\mathbf{e}_2 + x_3\mathbf{e}_3$
3. Show that if $\mathbf{f} = \nabla\phi$ then $\text{curl}\mathbf{f} = \mathbf{o}$.
4. Show that if $\mathbf{f} = \nabla\times\mathbf{a}$ then $\nabla\cdot\mathbf{f} = 0$.
5. Find the volume beneath the surface $x_1^2 + x_2^2 - x_3 = 0$ and above the square with vertices $(0,0)$, $(1,0)$, $(1,1)$ and $(0,1)$ in the $x_1 - x_2$ plane.
6. Find the Jacobian (and sketch lines of constant t_1, t_2) for the rotation
$$x_1 = t_1\cos\theta - t_2\sin\theta$$
$$x_2 = t_1\sin\theta + t_2\cos\theta$$
7. Find a unit normal to the circular cylinder with parametric representation
$$\mathbf{x}(t_1,t_2) = a\cos t_1\mathbf{e}_1 + a\sin t_1\mathbf{e}_2 + t_2\mathbf{e}_3, \quad 0\le t_1\le 2\pi, \quad 0\le t_1\le 1$$
8. Evaluate $\int_S \psi dS$ where $\psi = x_1 + x_2 + x_3$ and S is the plane surface $x_3 = x_1 + x_2$, $0\le x_2\le x_1$, $0\le x_1\le 1$.
9. Evaluate the flux integral $\int_S \mathbf{f}\cdot\mathbf{n}dS$ where $\mathbf{f} = \mathbf{e}_1 + 2\mathbf{e}_2 + 2\mathbf{e}_3$ and S is the cone $x_3 = a(x_1^2 + x_2^2)$, $x_3 \le a$ [Hint: first parameterise the surface with t_1, t_2.]
10. Prove the relations in (1.7.13). [Hint: first write the expressions in index notation.]
11. Use the divergence theorem to show that
$$\int_S \mathbf{x}\cdot\mathbf{n}dS = 3V,$$
where V is the volume enclosed by S (and $\mathbf{x}$ is the position vector).
12. Verify the divergence theorem for $\mathbf{v} = x_1^3\mathbf{e}_1 + x_2^3\mathbf{e}_2 + x_3^3\mathbf{e}_3$ where S is the surface of the sphere $x_1^2 + x_2^2 + x_3^2 = a^2$.
13. Interpret the divergence theorem (1.7.12) for the case when $\mathbf{v}$ is the velocity field. See (1.6.18, 1.7.8). Interpret also the case of $\text{div}\mathbf{v} = 0$.

14. Verify Stoke's theorem for $\mathbf{f} = x_2\mathbf{e}_1 + x_3\mathbf{e}_2 + x_1\mathbf{e}_3$ where *S* is $x_3 = 1 - x_1^2 - x_2^2 \geq 0$ (so that *C* is the circle of radius 1 in the $x_1 - x_2$ plane).
15. Verify Green's theorem for the case of $\psi_1 = x_1^2 - 2x_2, \psi_2 = x_1 + x_2$, with *C* the unit circle $x_1^2 + x_2^2 = 1$. The following relations might be useful:
$$\int_0^{2\pi} \sin^2\theta d\theta = \int_0^{2\pi} \cos^2\theta d\theta = \pi, \quad \int_0^{2\pi} \sin\theta\cos\theta d\theta = \int_0^{2\pi} \sin\theta\cos^2\theta d\theta = 0$$
16. Evaluate $\oint_C \mathbf{f} \cdot d\mathbf{x}$ using Green's theorem, where $\mathbf{f} = -x_2^3\mathbf{e}_1 + x_1^3\mathbf{e}_2$ and *C* is the circle $x_1^2 + x_2^2 = 4$.
17. Use Green's theorem to show that the double integral of the Laplacian of *p* over a region *R* is equivalent to the integral of $\partial p / \partial n = \text{grad} p \cdot \mathbf{n}$ around the curve *C* bounding the region:
$$\iint_R \nabla^2 p dx_1 dx_2 = \oint_C \frac{\partial p}{\partial n} ds$$
[Hint: Let $\psi_1 = -\partial p / \partial x_2, \psi_2 = +\partial p / \partial x_1$. Also, show that
$$\mathbf{n} = \frac{dx_2}{ds}\mathbf{e}_1 - \frac{dx_1}{ds}\mathbf{e}_2$$
is a unit normal to *C*, Fig. 1.7.10]

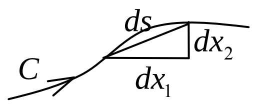


Figure 1.7.10: projection of a surface element onto a plane region

1.8 Tensors

Here the concept of the **tensor** is introduced. Tensors can be of different **orders** – zeroth-order tensors, first-order tensors, second-order tensors, and so on. Apart from the zeroth and first order tensors (see below), the second-order tensors are the most important tensors from a practical point of view, being important quantities in, amongst other topics, continuum mechanics, relativity, electromagnetism and quantum theory.

1.8.1 Zeroth and First Order Tensors

A **tensor of order zero** is simply another name for a scalar α.

A **first-order tensor** is simply another name for a vector **u**.

1.8.2 Second Order Tensors

Notation

Vectors:	lowercase bold-face Latin letters, e.g. **a**, **r**, **q**
2nd order Tensors:	uppercase bold-face Latin letters, e.g. **F**, **T**, **S**

Tensors as Linear Operators

A *second*-order tensor **T** may be *defined* as an operator that acts on a vector **u** generating another vector **v**, so that $\mathbf{T}(\mathbf{u}) = \mathbf{v}$, or[1]

$$\boxed{\mathbf{T}\cdot\mathbf{u} = \mathbf{v} \quad \text{or} \quad \mathbf{T}\mathbf{u} = \mathbf{v}} \qquad \textbf{Second-order Tensor} \qquad (1.8.1)$$

The second-order tensor **T** is a **linear operator** (or **linear transformation**)[2], which means that

$$\begin{aligned} \mathbf{T}(\mathbf{a}+\mathbf{b}) &= \mathbf{T}\mathbf{a} + \mathbf{T}\mathbf{b} && \ldots \text{ distributive} \\ \mathbf{T}(\alpha\mathbf{a}) &= \alpha(\mathbf{T}\mathbf{a}) && \ldots \text{ associative} \end{aligned}$$

This linearity can be viewed geometrically as in Fig. 1.8.1.

Note: the vector may also be defined in this way, as a mapping **u** that acts on a vector **v**, this time generating a scalar α, $\mathbf{u}\cdot\mathbf{v} = \alpha$. This transformation (the dot product) is linear (see properties (2,3) in §1.1.4). Thus a first-order tensor (vector) maps a first-order tensor into a zeroth-order tensor (scalar), whereas a second-order tensor maps a first-order tensor into a first-order tensor. It will be seen that a third-order tensor maps a first-order tensor into a second-order tensor, and so on.

---

[1] both these notations for the tensor operation are used; here, the convention of omitting the "dot" will be used

[2] An operator or transformation is a special function which maps elements of one type into elements of a similar type; here, vectors into vectors

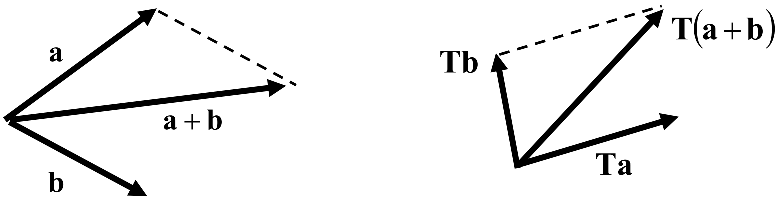


Figure 1.8.1: Linearity of the second order tensor

Further, two tensors **T** and **S** are said to be equal if and only if

$$\mathbf{Sv} = \mathbf{Tv}$$

for all vectors **v**.

Example (of a Tensor)

Suppose that **F** is an operator which transforms every vector into its mirror-image with respect to a given plane, Fig. 1.8.2. **F** transforms a vector into another vector and the transformation is linear, as can be seen geometrically from the figure. Thus **F** is a second-order tensor.

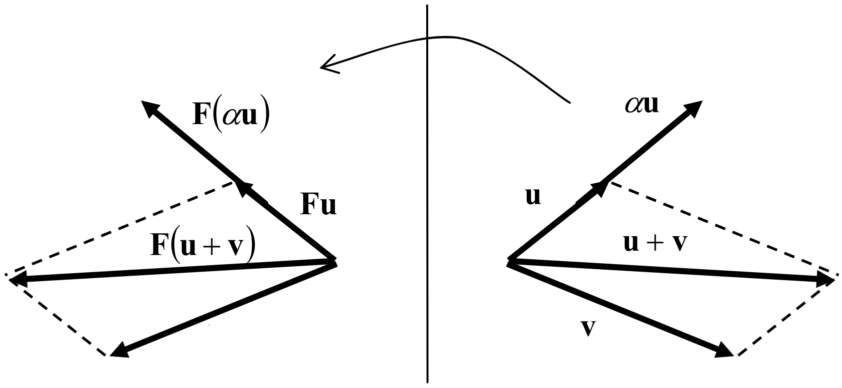


Figure 1.8.2: Mirror-imaging of vectors as a second order tensor mapping

■

Example (of a Tensor)

The combination $\mathbf{u}\times$ linearly transforms a vector into another vector and is thus a second-order tensor[3]. For example, consider a force **f** applied to a spanner at a distance **r** from the centre of the nut, Fig. 1.8.3. Then it can be said that the tensor $(\mathbf{r}\times)$ maps the force **f** into the (moment/torque) vector $\mathbf{r}\times\mathbf{f}$.

[3] Some authors use the notation $\tilde{\mathbf{u}}$ to denote $\mathbf{u}\times$

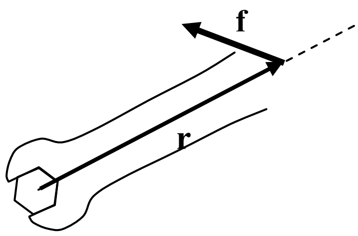


Figure 1.8.3: the force on a spanner

■

1.8.3 The Dyad (the tensor product)

The vector *dot product* and vector *cross product* have been considered in previous sections. A third vector product, the **tensor product** (or **dyadic product**), is important in the analysis of tensors of order 2 or more. The tensor product of two vectors **u** and **v** is written as[4]

$$\boxed{\mathbf{u} \otimes \mathbf{v}} \qquad \textbf{Tensor Product} \qquad (1.8.2)$$

This tensor product is itself a tensor of order two, and is called **dyad**:

$\mathbf{u} \cdot \mathbf{v}$	is a scalar	(a zeroth order tensor)
$\mathbf{u} \times \mathbf{v}$	is a vector	(a first order tensor)
$\mathbf{u} \otimes \mathbf{v}$	is a dyad	(a second order tensor)

It is best to *define* this dyad by what it *does*: it transforms a vector **w** into another vector with the direction of **u** according to the rule[5]

$$\boxed{(\mathbf{u} \otimes \mathbf{v})\mathbf{w} = \mathbf{u}(\mathbf{v} \cdot \mathbf{w})} \qquad \textbf{The Dyad Transformation} \qquad (1.8.3)$$

This relation defines the symbol "$\otimes$".

The length of the new vector is $|\mathbf{u}|$ times $\mathbf{v} \cdot \mathbf{w}$, and the new vector has the same direction as $\mathbf{u}$, Fig. 1.8.4. It can be seen that the dyad is a second order tensor, because it operates linearly on a vector to give another vector {▲Problem 2}.

Note that the dyad is not commutative, $\mathbf{u} \otimes \mathbf{v} \neq \mathbf{v} \otimes \mathbf{u}$. Indeed it can be seen clearly from the figure that $(\mathbf{u} \otimes \mathbf{v})\mathbf{w} \neq (\mathbf{v} \otimes \mathbf{u})\mathbf{w}$.

[4] many authors omit the $\otimes$ and write simply **uv**
[5] note that it is the two vectors that are beside each other (separated by a bracket) that get "dotted" together

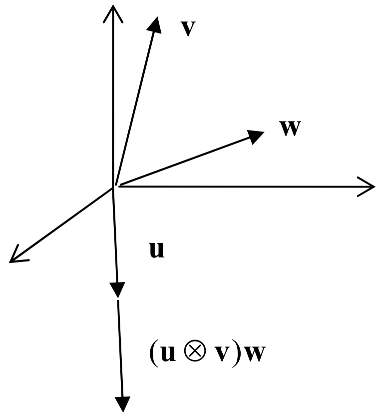


Figure 1.8.4: the dyad transformation

The following important relations follow from the above definition {▲Problem 4},

$$\begin{aligned}(\mathbf{u}\otimes\mathbf{v})(\mathbf{w}\otimes\mathbf{x}) &= (\mathbf{v}\cdot\mathbf{w})(\mathbf{u}\otimes\mathbf{x})\\ \mathbf{u}(\mathbf{v}\otimes\mathbf{w}) &= (\mathbf{u}\cdot\mathbf{v})\mathbf{w}\end{aligned} \tag{1.8.4}$$

It can be seen from these that the operation of the dyad on a vector is not commutative:

$$\mathbf{u}(\mathbf{v}\otimes\mathbf{w}) \neq (\mathbf{v}\otimes\mathbf{w})\mathbf{u} \tag{1.8.5}$$

Example (The Projection Tensor)

Consider the dyad $\mathbf{e}\otimes\mathbf{e}$. From the definition 1.8.3, $(\mathbf{e}\otimes\mathbf{e})\mathbf{u} = (\mathbf{e}\cdot\mathbf{u})\mathbf{e}$. But $\mathbf{e}\cdot\mathbf{u}$ is the projection of **u** onto a line through the unit vector **e**. Thus $(\mathbf{e}\cdot\mathbf{u})\mathbf{e}$ is the vector projection of **u** on **e**. For this reason $\mathbf{e}\otimes\mathbf{e}$ is called the **projection tensor**. It is usually denoted by **P**.

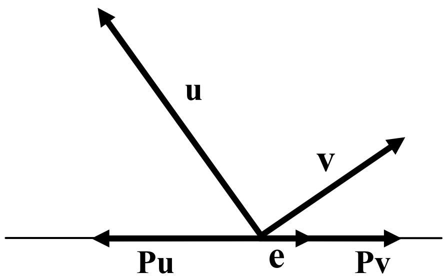


Figure 1.8.5: the projection tensor

■

1.8.4 Dyadics

A **dyadic** is a linear combination of these dyads (with scalar coefficients). An example might be

$$5(\mathbf{a} \otimes \mathbf{b}) + 3(\mathbf{c} \otimes \mathbf{d}) - 2(\mathbf{e} \otimes \mathbf{f})$$

This is clearly a second-order tensor. It will be seen in §1.9 that *every second-order tensor can be represented by a* dyadic, that is

$$\mathbf{T} = \alpha(\mathbf{a} \otimes \mathbf{b}) + \beta(\mathbf{c} \otimes \mathbf{d}) + \gamma(\mathbf{e} \otimes \mathbf{f}) + \cdots \tag{1.8.6}$$

Note: second-order tensors cannot, in general, be written as a dyad, $\mathbf{T} = \mathbf{a} \otimes \mathbf{b}$ – when they can, they are called **simple tensors**.

Example (Angular Momentum and the Moment of Inertia Tensor)

Suppose a rigid body is rotating so that every particle in the body is instantaneously moving in a circle about some axis fixed in space, Fig. 1.8.6.

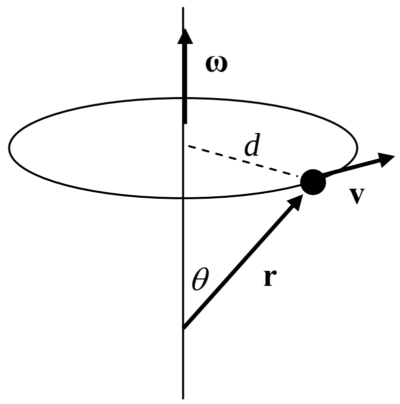


Figure 1.8.6: a particle in motion about an axis

The body's angular velocity $\boldsymbol{\omega}$ is defined as the vector whose magnitude is the angular speed ω and whose direction is along the axis of rotation. Then a particle's linear velocity is

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$$

where $v = wd$ is the linear speed, *d* is the distance between the axis and the particle, and **r** is the position vector of the particle from a fixed point *O* on the axis. The particle's **angular momentum** (or moment of momentum) **h** about the point *O* is defined to be

$$\mathbf{h} = m\mathbf{r} \times \mathbf{v}$$

where *m* is the mass of the particle. The angular momentum can be written as

$$\mathbf{h} = \hat{\mathbf{I}}\boldsymbol{\omega} \tag{1.8.8}$$

where $\hat{\mathbf{I}}$, a second-order tensor, is the **moment of inertia** of the particle about the point O, given by

$$\hat{\mathbf{I}} = m\left(|\mathbf{r}|^2 \mathbf{I} - \mathbf{r} \otimes \mathbf{r}\right) \tag{1.8.9}$$

where $\mathbf{I}$ is the identity tensor, i.e. $\mathbf{Ia} = \mathbf{a}$ for all vectors $\mathbf{a}$.

To show this, it must be shown that $\mathbf{r} \times \mathbf{v} = \left(|\mathbf{r}|^2 \mathbf{I} - \mathbf{r} \otimes \mathbf{r}\right)\boldsymbol{\omega}$. First examine $\mathbf{r} \times \mathbf{v}$. It is evidently a vector perpendicular to both $\mathbf{r}$ and $\mathbf{v}$ and in the plane of $\mathbf{r}$ and $\boldsymbol{\omega}$; its magnitude is

$$|\mathbf{r} \times \mathbf{v}| = |\mathbf{r}||\mathbf{v}| = |\mathbf{r}|^2 |\boldsymbol{\omega}| \sin\theta$$

Now (see Fig. 1.8.7)

$$\begin{aligned}\left(|\mathbf{r}|^2 \mathbf{I} - \mathbf{r} \otimes \mathbf{r}\right)\boldsymbol{\omega} &= |\mathbf{r}|^2 \boldsymbol{\omega} - \mathbf{r}(\mathbf{r} \cdot \boldsymbol{\omega}) \\ &= |\mathbf{r}|^2 |\boldsymbol{\omega}| \left(\mathbf{e}_\omega - \cos\theta \mathbf{e}_r\right)\end{aligned}$$

where $\mathbf{e}_\omega$ and $\mathbf{e}_r$ are unit vectors in the directions of $\boldsymbol{\omega}$ and $\mathbf{r}$ respectively. From the diagram, this is equal to $|\mathbf{r}|^2 |\boldsymbol{\omega}| \sin\theta \mathbf{e}_h$. Thus both expressions are equivalent, and one can indeed write $\mathbf{h} = \hat{\mathbf{I}}\boldsymbol{\omega}$ with $\hat{\mathbf{I}}$ defined by Eqn. 1.8.9: the second-order tensor $\hat{\mathbf{I}}$ maps the angular velocity vector $\boldsymbol{\omega}$ into the angular momentum vector $\mathbf{h}$ of the particle.

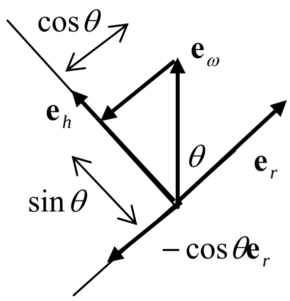


Figure 1.8.7: geometry of unit vectors for angular momentum calculation

■

1.8.5 The Vector Space of Second Order Tensors

The vector space of vectors and associated spaces were discussed in §1.2. Here, spaces of second order tensors are discussed.

As mentioned above, the second order tensor is a mapping on the vector space *V*,

$$\mathbf{T} : V \to V \tag{1.8.10}$$

and follows the rules

$$\begin{aligned} \mathbf{T}(\mathbf{a}+\mathbf{b}) &= \mathbf{Ta}+\mathbf{Tb} \\ \mathbf{T}(\alpha\mathbf{a}) &= \alpha(\mathbf{Ta}) \end{aligned} \tag{1.8.11}$$

for all $\mathbf{a}, \mathbf{b} \in V$ and $\alpha \in R$.

Denote the set of all second order tensors by V^2. Define then the sum of two tensors $\mathbf{S}, \mathbf{T} \in V^2$ through the relation

$$(\mathbf{S}+\mathbf{T})\mathbf{v} = \mathbf{Sv}+\mathbf{Tv} \tag{1.8.12}$$

and the product of a scalar $\alpha \in R$ and a tensor $\mathbf{T} \in V^2$ through

$$(\alpha\mathbf{T})\mathbf{v} = \alpha\mathbf{Tv} \tag{1.8.13}$$

Define an identity tensor $\mathbf{I} \in V^2$ through

$$\mathbf{Iv} = \mathbf{v}, \quad \text{for all } \mathbf{v} \in V \tag{1.8.14}$$

and a zero tensor $\mathbf{O} \in V^2$ through

$$\mathbf{Ov} = \mathbf{o}, \quad \text{for all } \mathbf{v} \in V \tag{1.8.15}$$

It follows from the definition 1.8.11 that V^2 has the structure of a real vector space, that is, the sum $\mathbf{S}+\mathbf{T} \in V^2$, the product $\alpha\mathbf{T} \in V^2$, and the following 8 axioms hold:

1. for any $\mathbf{A}, \mathbf{B}, \mathbf{C} \in V^2$, one has $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$
2. there exists an element $\mathbf{O} \in V^2$ such that $\mathbf{T}+\mathbf{O} = \mathbf{O}+\mathbf{T} = \mathbf{T}$ for every $\mathbf{T} \in V^2$
3. for each $\mathbf{T} \in V^2$ there exists an element $-\mathbf{T} \in V^2$, called the negative of **T**, such that $\mathbf{T}+(-\mathbf{T}) = (-\mathbf{T})+\mathbf{T} = 0$
4. for any $\mathbf{S}, \mathbf{T} \in V^2$, one has $\mathbf{S}+\mathbf{T} = \mathbf{T}+\mathbf{S}$
5. for any $\mathbf{S}, \mathbf{T} \in V^2$ and scalar $\alpha \in R$, $\alpha(\mathbf{S}+\mathbf{T}) = \alpha\mathbf{S}+\alpha\mathbf{T}$
6. for any $\mathbf{T} \in V^2$ and scalars $\alpha, \beta \in R$, $(\alpha+\beta)\mathbf{T} = \alpha\mathbf{T}+\beta\mathbf{T}$
7. for any $\mathbf{T} \in V^2$ and scalars $\alpha, \beta \in R$, $\alpha(\beta\mathbf{T}) = (\alpha\beta)\mathbf{T}$
8. for the unit scalar $1 \in R$, $1\mathbf{T} = \mathbf{T}$ for any $\mathbf{T} \in V^2$.

1.8.6 Problems

1. Consider the function **f** which transforms a vector **v** into $\mathbf{a}\cdot\mathbf{v}+\beta$. Is **f** a tensor (of order one)? [Hint: test to see whether the transformation is linear, by examining $\mathbf{f}(\alpha\mathbf{u}+\mathbf{v})$.]
2. Show that the dyad is a linear operator, in other words, show that $(\mathbf{u}\otimes\mathbf{v})(\alpha\mathbf{w}+\beta\mathbf{x})=\alpha(\mathbf{u}\otimes\mathbf{v})\mathbf{w}+\beta(\mathbf{u}\otimes\mathbf{v})\mathbf{x}$
3. When is $\mathbf{a}\otimes\mathbf{b}=\mathbf{b}\otimes\mathbf{a}$?
4. Prove that
   (i) $(\mathbf{u}\otimes\mathbf{v})(\mathbf{w}\otimes\mathbf{x})=(\mathbf{v}\cdot\mathbf{w})(\mathbf{u}\otimes\mathbf{x})$ [Hint: post-“multiply” both sides of the definition (1.8.3) by $\otimes\mathbf{x}$; then show that $((\mathbf{u}\otimes\mathbf{v})\mathbf{w})\otimes\mathbf{x}=(\mathbf{u}\otimes\mathbf{v})(\mathbf{w}\otimes\mathbf{x})$.]
   (ii) $\mathbf{u}(\mathbf{v}\otimes\mathbf{w})=(\mathbf{u}\cdot\mathbf{v})\mathbf{w}$ [hint: pre “multiply” both sides by $\mathbf{x}\otimes$ and use the result of (i)]
5. Consider the dyadic (tensor) $\mathbf{a}\otimes\mathbf{a}+\mathbf{b}\otimes\mathbf{b}$. Show that this tensor orthogonally projects every vector **v** onto the plane formed by **a** and **b** (sketch a diagram).
6. Draw a sketch to show the meaning of $\mathbf{u}\cdot(\mathbf{P}\mathbf{v})$, where **P** is the projection tensor. What is the order of the resulting tensor?
7. Prove that $\mathbf{a}\otimes\mathbf{b}-\mathbf{b}\otimes\mathbf{a}=(\mathbf{b}\times\mathbf{a})\times$.

1.9 Cartesian Tensors

As with the vector, a (higher order) tensor is a mathematical object which represents many physical phenomena and which exists independently of any coordinate system. In what follows, a Cartesian coordinate system is used to describe tensors.

1.9.1 Cartesian Tensors

A second order tensor and the vector it operates on can be described in terms of Cartesian components. For example, $(\mathbf{a} \otimes \mathbf{b})\mathbf{c}$, with $\mathbf{a} = 2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$, $\mathbf{b} = \mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3$ and $\mathbf{c} = -\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, is

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{c} = \mathbf{a}(\mathbf{b} \cdot \mathbf{c}) = 4\mathbf{e}_1 + 2\mathbf{e}_2 - 2\mathbf{e}_3$$

Example (The Unit Dyadic or Identity Tensor)

The **identity tensor**, or **unit tensor**, **I**, which maps every vector onto itself, has been introduced in the previous section. The Cartesian representation of **I** is

$$\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 \equiv \mathbf{e}_i \otimes \mathbf{e}_i \qquad (1.9.1)$$

This follows from

$$\begin{aligned}
(\mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3)\mathbf{u} &= (\mathbf{e}_1 \otimes \mathbf{e}_1)\mathbf{u} + (\mathbf{e}_2 \otimes \mathbf{e}_2)\mathbf{u} + (\mathbf{e}_3 \otimes \mathbf{e}_3)\mathbf{u} \\
&= \mathbf{e}_1(\mathbf{e}_1 \cdot \mathbf{u}) + \mathbf{e}_2(\mathbf{e}_2 \cdot \mathbf{u}) + \mathbf{e}_3(\mathbf{e}_3 \cdot \mathbf{u}) \\
&= u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3 \\
&= \mathbf{u}
\end{aligned}$$

Note also that the identity tensor can be written as $\mathbf{I} = \delta_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)$, in other words the Kronecker delta gives the components of the identity tensor in a Cartesian coordinate system.

■

Second Order Tensor as a Dyadic

In what follows, it will be shown that a second order tensor can always be written as a dyadic involving the Cartesian base vectors $\mathbf{e}_i$ [1].

Consider an arbitrary second-order tensor **T** which operates on **a** to produce **b**, $\mathbf{T}(\mathbf{a}) = \mathbf{b}$, or $\mathbf{T}(a_i\mathbf{e}_i) = \mathbf{b}$. From the linearity of **T**,

[1] this can be generalised to the case of non-Cartesian base vectors, which might not be orthogonal nor of unit magnitude (see §1.16)

$$a_1\mathbf{T}(\mathbf{e}_1) + a_2\mathbf{T}(\mathbf{e}_2) + a_3\mathbf{T}(\mathbf{e}_3) = \mathbf{b}$$

Just as **T** transforms **a** into **b**, it transforms the base vectors $\mathbf{e}_i$ into some other vectors; suppose that $\mathbf{T}(\mathbf{e}_1) = \mathbf{u}$, $\mathbf{T}(\mathbf{e}_2) = \mathbf{v}$, $\mathbf{T}(\mathbf{e}_3) = \mathbf{w}$, then

$$\begin{aligned}\mathbf{b} &= a_1\mathbf{u} + a_2\mathbf{v} + a_3\mathbf{w} \\ &= (\mathbf{a}\cdot\mathbf{e}_1)\mathbf{u} + (\mathbf{a}\cdot\mathbf{e}_2)\mathbf{v} + (\mathbf{a}\cdot\mathbf{e}_3)\mathbf{w} \\ &= (\mathbf{u}\otimes\mathbf{e}_1)\mathbf{a} + (\mathbf{v}\otimes\mathbf{e}_2)\mathbf{a} + (\mathbf{w}\otimes\mathbf{e}_3)\mathbf{a} \\ &= [\mathbf{u}\otimes\mathbf{e}_1 + \mathbf{v}\otimes\mathbf{e}_2 + \mathbf{w}\otimes\mathbf{e}_3]\mathbf{a}\end{aligned}$$

and so

$$\mathbf{T} = \mathbf{u}\otimes\mathbf{e}_1 + \mathbf{v}\otimes\mathbf{e}_2 + \mathbf{w}\otimes\mathbf{e}_3 \tag{1.9.2}$$

which is indeed a dyadic.

Cartesian components of a Second Order Tensor

The second order tensor **T** can be written in terms of components and base vectors as follows: write the vectors **u**, **v** and **w** in (1.9.2) in component form, so that

$$\begin{aligned}\mathbf{T} &= (u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3)\otimes\mathbf{e}_1 + (\cdots)\otimes\mathbf{e}_2 + (\cdots)\otimes\mathbf{e}_3 \\ &= u_1\mathbf{e}_1\otimes\mathbf{e}_1 + u_2\mathbf{e}_2\otimes\mathbf{e}_1 + u_3\mathbf{e}_3\otimes\mathbf{e}_1 + \cdots\end{aligned}$$

Introduce nine scalars T_{ij} by letting $u_i = T_{i1}$, $v_i = T_{i2}$, $w_i = T_{i3}$, so that

$$\boxed{\begin{aligned}\mathbf{T} = {} & T_{11}\mathbf{e}_1\otimes\mathbf{e}_1 + T_{12}\mathbf{e}_1\otimes\mathbf{e}_2 + T_{13}\mathbf{e}_1\otimes\mathbf{e}_3 \\ & + T_{21}\mathbf{e}_2\otimes\mathbf{e}_1 + T_{22}\mathbf{e}_2\otimes\mathbf{e}_2 + T_{23}\mathbf{e}_2\otimes\mathbf{e}_3 \\ & + T_{31}\mathbf{e}_3\otimes\mathbf{e}_1 + T_{32}\mathbf{e}_3\otimes\mathbf{e}_2 + T_{33}\mathbf{e}_3\otimes\mathbf{e}_3\end{aligned}} \quad \textbf{Second-order Cartesian Tensor} \tag{1.9.3}$$

These nine scalars T_{ij} are the components of the second order tensor **T** in the Cartesian coordinate system. In index notation,

$$\mathbf{T} = T_{ij}(\mathbf{e}_i\otimes\mathbf{e}_j)$$

Thus whereas a vector has three components, a second order tensor has *nine* components. Similarly, whereas the three vectors $\{\mathbf{e}_i\}$ form a basis for the space of vectors, the nine dyads $\{\mathbf{e}_i\otimes\mathbf{e}_j\}$ form a basis for the space of tensors, i.e. all second order tensors can be expressed as a linear combination of these basis tensors.

It can be shown that the components of a second-order tensor can be obtained directly from {▲Problem 1}

$$\boxed{T_{ij} = \mathbf{e}_i\mathbf{T}\mathbf{e}_j} \qquad \textbf{Components of a Tensor} \tag{1.9.4}$$

which is the tensor expression analogous to the vector expression $u_i = \mathbf{e}_i \cdot \mathbf{u}$. Note that, in Eqn. 1.9.4, the components can be written simply as $\mathbf{e}_i \mathbf{T} \mathbf{e}_j$ (without a "dot"), since $\mathbf{e}_i \cdot \mathbf{T}\mathbf{e}_j = \mathbf{e}_i \mathbf{T} \cdot \mathbf{e}_j$.

Example (The Stress Tensor)

Define the traction vector **t** acting on a surface element within a material to be the force acting on that element[2] divided by the area of the element, Fig. 1.9.1. Let **n** be a vector normal to the surface. The **stress** $\boldsymbol{\sigma}$ is defined to be that second order tensor which maps **n** onto **t**, according to

$$\boxed{\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}} \qquad \textbf{The Stress Tensor} \qquad (1.9.5)$$

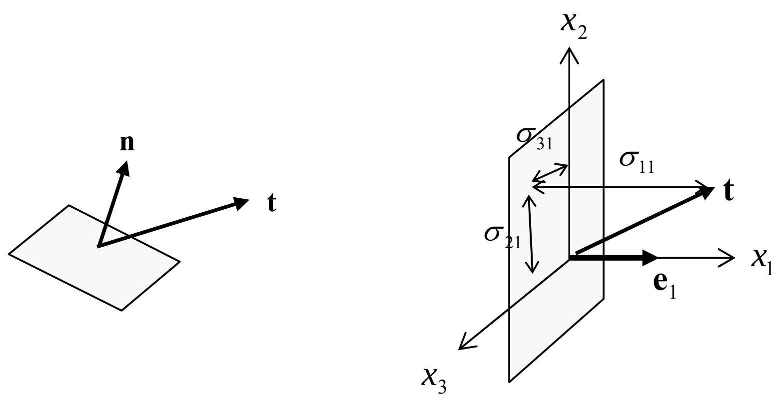


Figure 1.9.1: stress acting on a plane

If one now considers a coordinate system with base vectors $\mathbf{e}_i$, then $\boldsymbol{\sigma} = \sigma_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$ and, for example,

$$\boldsymbol{\sigma}\mathbf{e}_1 = \sigma_{11}\mathbf{e}_1 + \sigma_{21}\mathbf{e}_2 + \sigma_{31}\mathbf{e}_3$$

Thus the components σ_{11}, σ_{21} and σ_{31} of the stress tensor are the three components of the traction vector which acts on the plane with normal $\mathbf{e}_1$.

Augustin-Louis Cauchy was the first to regard stress as a linear map of the normal vector onto the traction vector; hence the name "tensor", from the French for stress, *tension*.

■

[2] this force would be due, for example, to intermolecular forces within the material: the particles on one side of the surface element exert a force on the particles on the other side

Higher Order Tensors

The above can be generalised to tensors of order three and higher. The following notation will be used:

α, β, γ	…	0th-order tensors	("scalars")
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	…	1st-order tensors	("vectors")
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	…	2nd-order tensors	("dyadics")
$\boldsymbol{\mathsf{A}}, \boldsymbol{\mathsf{B}}, \boldsymbol{\mathsf{C}}$	…	3rd-order tensors	("triadics")
$\mathsf{A}, \mathsf{B}, \mathsf{C}$	…	4th-order tensors	("tetradics")

An important third-order tensor is the **permutation tensor**, defined by

$$\boldsymbol{\mathsf{E}} = \varepsilon_{ijk} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \tag{1.9.6}$$

whose components are those of the permutation symbol, Eqns. 1.3.10-1.3.13.

A fourth-order tensor can be written as

$$\mathsf{A} = A_{ijkl} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l \tag{1.9.7}$$

It can be seen that a zeroth-order tensor (scalar) has $3^0 = 1$ component, a first-order tensor has $3^1 = 3$ components, a second-order tensor has $3^2 = 9$ components, so $\boldsymbol{\mathsf{A}}$ has $3^3 = 27$ components and A has 81 components.

1.9.2 Simple Contraction

Tensor/vector operations can be written in component form, for example,

$$\begin{aligned} \mathbf{Ta} &= T_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j) a_k \mathbf{e}_k \\ &= T_{ij} a_k [(\mathbf{e}_i \otimes \mathbf{e}_j) \mathbf{e}_k] \\ &= T_{ij} a_k \delta_{jk} \mathbf{e}_i \\ &= T_{ij} a_j \mathbf{e}_i \end{aligned} \tag{1.9.8}$$

This operation is called **simple contraction**, because the order of the tensors is contracted – to begin there was a tensor of order 2 and a tensor of order 1, and to end there is a tensor of order 1 (it is called "simple" to distinguish it from "double" contraction – see below). This is always the case – when a tensor operates on another in this way, the order of the result will be *two* less than the sum of the original orders.

An example of simple contraction of two second order tensors has already been seen in Eqn. 1.8.4a; the tensors there were simple tensors (dyads). Here is another example:

$$\begin{aligned}\mathbf{TS} &= T_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)S_{kl}(\mathbf{e}_k \otimes \mathbf{e}_l) \\ &= T_{ij}S_{kl}\left[(\mathbf{e}_i \otimes \mathbf{e}_j)(\mathbf{e}_k \otimes \mathbf{e}_l)\right] \\ &= T_{ij}S_{kl}\delta_{jk}(\mathbf{e}_i \otimes \mathbf{e}_l) \\ &= T_{ij}S_{jl}(\mathbf{e}_i \otimes \mathbf{e}_l)\end{aligned} \tag{1.9.9}$$

From the above, the simple contraction of two second order tensors results in another second order tensor. If one writes $\mathbf{A} = \mathbf{TS}$, then the components of the new tensor are related to those of the original tensors through $A_{ij} = T_{ik}S_{kj}$.

Note that, in general,

$$\begin{aligned} \mathbf{AB} &\neq \mathbf{BA} \\ (\mathbf{AB})\mathbf{C} &= \mathbf{A}(\mathbf{BC}) && \text{… associative} \\ \mathbf{A}(\mathbf{B}+\mathbf{C}) &= \mathbf{AB}+\mathbf{AC} && \text{… distributive}\end{aligned} \tag{1.9.10}$$

The associative and distributive properties follow from the fact that a tensor is by definition a linear operator, §1.8.2; they apply to tensors of any order, for example,

$$(\mathbf{AB})\mathbf{v} = \mathbf{A}(\mathbf{Bv}) \tag{1.9.11}$$

To deal with tensors of any order, all one has to remember is how simple tensors operate on each other – the two vectors which are beside each other are the ones which are "dotted" together:

$$\begin{aligned}(\mathbf{a}\otimes\mathbf{b})\mathbf{c} &= (\mathbf{b}\cdot\mathbf{c})\mathbf{a} \\ (\mathbf{a}\otimes\mathbf{b})(\mathbf{c}\otimes\mathbf{d}) &= (\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\otimes\mathbf{d}) \\ (\mathbf{a}\otimes\mathbf{b})(\mathbf{c}\otimes\mathbf{d}\otimes\mathbf{e}) &= (\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\otimes\mathbf{d}\otimes\mathbf{e}) \\ (\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c})(\mathbf{d}\otimes\mathbf{e}\otimes\mathbf{f}) &= (\mathbf{c}\cdot\mathbf{d})(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{e}\otimes\mathbf{f})\end{aligned} \tag{1.9.12}$$

An example involving a higher order tensor is

$$\begin{aligned}\boldsymbol{\mathsf{A}}\cdot\mathbf{E} &= A_{ijkl}E_{mn}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\otimes\mathbf{e}_l)(\mathbf{e}_m\otimes\mathbf{e}_n) \\ &= A_{ijkl}E_{ln}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\otimes\mathbf{e}_n)\end{aligned}$$

and

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= \alpha \\ \mathbf{AB} &= \mathbf{C} \\ \mathbf{Au} &= \mathbf{v} \\ \boldsymbol{\mathsf{A}}\mathbf{b} &= \mathbf{C} \\ \mathbf{AB} &= \boldsymbol{\mathsf{C}}\end{aligned}$$

Note the relation (analogous to the vector relation $\mathbf{a}(\mathbf{b}\otimes\mathbf{c})\mathbf{d} = (\mathbf{a}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{d})$, which follows directly from the dyad definition 1.8.3) {▲Problem 10}

$$\mathbf{A}(\mathbf{B}\otimes\mathbf{C})\mathbf{D} = (\mathbf{AB})\otimes(\mathbf{CD}) \tag{1.9.13}$$

Powers of Tensors

Integral powers of tensors are defined inductively by $\mathbf{T}^0 = \mathbf{I}$, $\mathbf{T}^n = \mathbf{T}^{n-1}\mathbf{T}$, so, for example,

$$\boxed{\mathbf{T}^2 = \mathbf{TT}} \qquad \textbf{The Square of a Tensor} \tag{1.9.14}$$

$\mathbf{T}^3 = \mathbf{TTT}$, etc.

1.9.3 Double Contraction

Double contraction, as the name implies, contracts the tensors twice as much a simple contraction. Thus, where the sum of the orders of two tensors is reduced by two in the simple contraction, the sum of the orders is reduced by four in double contraction. The double contraction is denoted by a colon (:), e.g. $\mathbf{T}:\mathbf{S}$.

First, define the double contraction of simple tensors (dyads) through

$$(\mathbf{a}\otimes\mathbf{b}):(\mathbf{c}\otimes\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) \tag{1.9.15}$$

So in double contraction, one takes the scalar product of four vectors which are adjacent to each other, according to the following rule:

$$(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}):(\mathbf{d}\otimes\mathbf{e}\otimes\mathbf{f}) = (\mathbf{b}\cdot\mathbf{d})(\mathbf{c}\cdot\mathbf{e})(\mathbf{a}\otimes\mathbf{f})$$

For example,

$$\begin{aligned}\mathbf{T}:\mathbf{S} &= T_{ij}(\mathbf{e}_i\otimes\mathbf{e}_j):S_{kl}(\mathbf{e}_k\otimes\mathbf{e}_l)\\ &= T_{ij}S_{kl}\left[(\mathbf{e}_i\cdot\mathbf{e}_k)(\mathbf{e}_j\cdot\mathbf{e}_l)\right]\\ &= T_{ij}S_{ij}\end{aligned} \tag{1.9.16}$$

which is, as expected, a scalar.

Here is another example, the contraction of the two second order tensors **I** (see Eqn. 1.9.1) and $\mathbf{u}\otimes\mathbf{v}$,

$$\begin{aligned}\mathbf{I}:\mathbf{u}\otimes\mathbf{v} &= (\mathbf{e}_i\otimes\mathbf{e}_i):(\mathbf{u}\otimes\mathbf{v})\\ &= (\mathbf{e}_i\cdot\mathbf{u})(\mathbf{e}_i\cdot\mathbf{v})\\ &= u_i v_i\\ &= \mathbf{u}\cdot\mathbf{v}\end{aligned} \tag{1.9.17}$$

so that the scalar product of two vectors can be written in the form of a double contraction involving the Identity Tensor.

An example of double contraction involving the permutation tensor 1.9.6 is {▲Problem 11}

$$\mathbf{E} : (\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \times \mathbf{v} \tag{1.9.18}$$

It can be shown that the components of a fourth order tensor are given by (compare with Eqn. 1.9.4)

$$A_{ijkl} = (\mathbf{e}_i \otimes \mathbf{e}_j) : \mathbf{A} : (\mathbf{e}_k \otimes \mathbf{e}_l) \tag{1.9.19}$$

In summary then,

$$\begin{aligned} \mathbf{A} : \mathbf{B} &= \beta \\ \mathbf{A} : \mathbf{b} &= \gamma \\ \mathbf{A} : \mathbf{B} &= \mathbf{c} \\ \mathbf{A} : \mathbf{B} &= \mathbf{C} \end{aligned}$$

Note the following identities:

$$\begin{aligned} (\mathbf{A} \otimes \mathbf{B}) : \mathbf{C} &= \mathbf{A}(\mathbf{B} : \mathbf{C}) = (\mathbf{B} : \mathbf{C})\mathbf{A} \\ \mathbf{A} : (\mathbf{B} \otimes \mathbf{C}) &= \mathbf{C}(\mathbf{A} : \mathbf{B}) = (\mathbf{A} : \mathbf{B})\mathbf{C} \\ (\mathbf{A} \otimes \mathbf{B}) : (\mathbf{C} \otimes \mathbf{D}) &= (\mathbf{B} : \mathbf{C})(\mathbf{A} \otimes \mathbf{D}) = (\mathbf{A} \otimes \mathbf{D})(\mathbf{B} : \mathbf{C}) \end{aligned} \tag{1.9.20}$$

Note: There are many operations that can be defined and performed with tensors. The two most important operations, the ones which arise most in practice, are the simple and double contractions defined above. Other possibilities are:

(a) double contraction with two "horizontal" dots, $\mathbf{T} \cdot\cdot \mathbf{S}$, $\mathbf{A} \cdot\cdot \mathbf{b}$, etc., which is based on the definition of the following operation as applied to simple tensors:

$$(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}) \cdot\cdot (\mathbf{d} \otimes \mathbf{e} \otimes \mathbf{f}) \equiv (\mathbf{b} \cdot \mathbf{e})(\mathbf{c} \cdot \mathbf{d})(\mathbf{a} \otimes \mathbf{f})$$

(b) operations involving one cross $(\times)$: $(\mathbf{a} \otimes \mathbf{b}) \times (\mathbf{c} \otimes \mathbf{d}) \equiv (\mathbf{a} \otimes \mathbf{d}) \otimes (\mathbf{b} \times \mathbf{c})$

(c) "double" operations involving the cross $(\times)$ and dot:

$$(\mathbf{a} \otimes \mathbf{b}) \overset{\times}{\underset{\times}{}} (\mathbf{c} \otimes \mathbf{d}) \equiv (\mathbf{a} \times \mathbf{c}) \otimes (\mathbf{b} \times \mathbf{d})$$

$$(\mathbf{a} \otimes \mathbf{b}) \overset{\times}{\underset{\cdot}{}} (\mathbf{c} \otimes \mathbf{d}) \equiv (\mathbf{a} \times \mathbf{c})(\mathbf{b} \cdot \mathbf{d})$$

$$(\mathbf{a} \otimes \mathbf{b}) \overset{\cdot}{\underset{\times}{}} (\mathbf{c} \otimes \mathbf{d}) \equiv (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \times \mathbf{d})$$

1.9.4 Index Notation

The index notation for single and double contraction of tensors of any order can easily be remembered. From the above, a single contraction of two tensors implies that the indices "beside each other" are the same[3], and a double contraction implies that a pair of indices is repeated. Thus, for example, in both symbolic and index notation:

$$\begin{aligned} \mathbf{AB} = \mathbf{C} &\qquad A_{ijm}B_{mk} = C_{ijk} \\ \mathbf{A} : \mathbf{B} = \mathbf{c} &\qquad A_{ijk}B_{jk} = c_i \end{aligned} \tag{1.9.21}$$

1.9.5 Matrix Notation

Here the matrix notation of §1.4 is extended to include second-order tensors[4]. The Cartesian components of a second-order tensor can conveniently be written as a 3×3 matrix,

$$[\mathbf{T}] = \begin{bmatrix} \mathrm{T}_{11} & \mathrm{T}_{12} & \mathrm{T}_{13} \\ \mathrm{T}_{21} & \mathrm{T}_{22} & \mathrm{T}_{13} \\ \mathrm{T}_{31} & \mathrm{T}_{32} & \mathrm{T}_{33} \end{bmatrix}$$

The operations involving vectors and second-order tensors can now be written in terms of matrices, for example,

$$\mathbf{Tu} = [\mathbf{T}][\mathbf{u}] = \begin{bmatrix} \mathrm{T}_{11} & \mathrm{T}_{12} & \mathrm{T}_{13} \\ \mathrm{T}_{21} & \mathrm{T}_{22} & \mathrm{T}_{13} \\ \mathrm{T}_{31} & \mathrm{T}_{32} & \mathrm{T}_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} T_{11}u_1 + T_{12}u_2 + T_{13}u_3 \\ T_{21}u_1 + T_{22}u_2 + T_{23}u_3 \\ T_{31}u_1 + T_{32}u_2 + T_{33}u_3 \end{bmatrix}$$

symbolic notation

"short" matrix notation

"full" matrix notation

The tensor product can be written as (see §1.4.1)

$$\mathbf{u} \otimes \mathbf{v} = [\mathbf{u}][\mathbf{v}^{\mathrm{T}}] = \begin{bmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{bmatrix} \tag{1.9.22}$$

which is consistent with the definition of the dyadic transformation, Eqn. 1.8.3.

[3] compare with the "beside each other rule" for matrix multiplication given in §1.4.1
[4] the matrix notation cannot be used for higher-order tensors

1.9.6 Problems

1. Use Eqn. 1.9.3 to show that the component T_{11} of a tensor $\mathbf{T}$ can be evaluated from $\mathbf{e}_1\mathbf{T}\mathbf{e}_1$, and that $T_{12} = \mathbf{e}_1\mathbf{T}\mathbf{e}_2$ (and so on, so that $T_{ij} = \mathbf{e}_i\mathbf{T}\mathbf{e}_j$).
2. Evaluate $\mathbf{aT}$ using the index notation (for a Cartesian basis). What is this operation called? Is your result equal to $\mathbf{Ta}$, in other words is this operation commutative? Now carry out this operation for two vectors, i.e. $\mathbf{a} \cdot \mathbf{b}$. Is it commutative in this case?
3. Evaluate the simple contractions $\mathbf{Ab}$ and $\mathbf{AB}$, with respect to a Cartesian coordinate system (use index notation).
4. Evaluate the double contraction $\mathbf{A} : \mathbf{B}$ (use index notation).
5. Show that, using a Cartesian coordinate system and the index notation, that the double contraction $\mathbf{A} : \mathbf{b}$ is a scalar. Write this scalar out in full in terms of the components of $\mathbf{A}$ and $\mathbf{b}$.
6. Consider the second-order tensors
$$\mathbf{D} = 3\mathbf{e}_1 \otimes \mathbf{e}_1 + 2\mathbf{e}_2 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_3 + 5\mathbf{e}_3 \otimes \mathbf{e}_3$$
$$\mathbf{F} = 4\mathbf{e}_1 \otimes \mathbf{e}_3 + 6\mathbf{e}_2 \otimes \mathbf{e}_2 - 3\mathbf{e}_3 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3$$
Compute $\mathbf{DF}$ and $\mathbf{F} : \mathbf{D}$.
7. Consider the second-order tensor
$$\mathbf{D} = 3\mathbf{e}_1 \otimes \mathbf{e}_1 - 4\mathbf{e}_1 \otimes \mathbf{e}_2 + 2\mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 .$$
Determine the image of the vector $\mathbf{r} = 4\mathbf{e}_1 + 2\mathbf{e}_2 + 5\mathbf{e}_3$ when $\mathbf{D}$ operates on it.
8. Write the following out in full – are these the components of scalars, vectors or second order tensors?
   (a) B_{ii}
   (b) C_{kkj}
   (c) B_{mn}
   (d) $a_i b_j A_{ij}$
9. Write $(\mathbf{a} \otimes \mathbf{b}) : (\mathbf{c} \otimes \mathbf{d})$ in terms of the components of the four vectors. What is the order of the resulting tensor?
10. Verify Eqn. 1.9.13.
11. Show that $\mathbf{E} : (\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \times \mathbf{v}$ – see (1.9.6, 1.9.18). [Hint: use the definition of the cross product in terms of the permutation symbol, (1.3.14), and the fact that $\varepsilon_{ijk} = -\varepsilon_{kji}$.]

1.10 Special Second Order Tensors & Properties of Second Order Tensors

In this section will be examined a number of special second order tensors, and special properties of second order tensors, which play important roles in tensor analysis. Many of the concepts will be familiar from Linear Algebra and Matrices. The following will be discussed:

- The Identity tensor
- Transpose of a tensor
- Trace of a tensor
- Norm of a tensor
- Determinant of a tensor
- Inverse of a tensor
- Orthogonal tensors
- Rotation Tensors
- Change of Basis Tensors
- Symmetric and Skew-symmetric tensors
- Axial vectors
- Spherical and Deviatoric tensors
- Positive Definite tensors

1.10.1 The Identity Tensor

The linear transformation which transforms every tensor into itself is called the **identity tensor**. This special tensor is denoted by **I** so that, for example,

$$\mathbf{Ia} = \mathbf{a} \quad \text{for any vector a}$$

In particular, $\mathbf{Ie}_1 = \mathbf{e}_1, \mathbf{Ie}_2 = \mathbf{e}_2, \mathbf{Ie}_3 = \mathbf{e}_3$, from which it follows that, for a Cartesian coordinate system, $I_{ij} = \delta_{ij}$. In matrix form,

$$[\mathbf{I}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.10.1}$$

1.10.2 The Transpose of a Tensor

The **transpose** of a second order tensor **A** with components A_{ij} is the tensor $\mathbf{A}^{\mathrm{T}}$ with components A_{ji}; so the transpose swaps the indices,

$$\boxed{\mathbf{A} = A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \quad \mathbf{A}^{\mathrm{T}} = A_{ji}\mathbf{e}_i \otimes \mathbf{e}_j} \quad \textbf{Transpose of a Second-Order Tensor} \tag{1.10.2}$$

In matrix notation,

$$[\mathbf{A}] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \qquad [\mathbf{A}^{\mathrm{T}}] = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

Some useful properties and relations involving the transpose are {▲Problem 2}:

$$\begin{aligned} \left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} &= \mathbf{A} \\ (\alpha\mathbf{A} + \beta\mathbf{B})^{\mathrm{T}} &= \alpha\mathbf{A}^{\mathrm{T}} + \beta\mathbf{B}^{\mathrm{T}} \\ (\mathbf{u} \otimes \mathbf{v})^{\mathrm{T}} &= \mathbf{v} \otimes \mathbf{u} \\ \mathbf{Tu} = \mathbf{uT}^{\mathrm{T}}, \quad \mathbf{uT} &= \mathbf{T}^{\mathrm{T}}\mathbf{u} \\ (\mathbf{AB})^{\mathrm{T}} &= \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \\ \mathbf{A} : \mathbf{B} &= \mathbf{A}^{\mathrm{T}} : \mathbf{B}^{\mathrm{T}} \\ (\mathbf{u} \otimes \mathbf{v})\mathbf{A} &= \mathbf{u} \otimes (\mathbf{A}^{\mathrm{T}}\mathbf{v}) \\ \mathbf{A} : (\mathbf{BC}) = \left(\mathbf{B}^{\mathrm{T}}\mathbf{A}\right) &: \mathbf{C} = \left(\mathbf{AC}^{\mathrm{T}}\right) : \mathbf{B} \end{aligned} \tag{1.10.3}$$

A formal definition of the transpose which does not rely on any particular coordinate system is as follows: the transpose of a second-order tensor is that tensor which satisfies the identity[1]

$$\mathbf{u} \cdot \mathbf{Av} = \mathbf{v} \cdot \mathbf{A}^{\mathrm{T}}\mathbf{u} \tag{1.10.4}$$

for all vectors **u** and **v**. To see that Eqn. 1.10.4 implies 1.10.2, first note that, for the present purposes, a convenient way of writing the components A_{ij} of the second-order tensor **A** is $(\mathbf{A})_{ij}$. From Eqn. 1.9.4, $(\mathbf{A})_{ij} = \mathbf{e}_i \cdot \mathbf{Ae}_j$ and the components of the transpose can be written as $\left(\mathbf{A}^{\mathrm{T}}\right)_{ij} = \mathbf{e}_i \cdot \mathbf{A}^{\mathrm{T}}\mathbf{e}_j$. Then, from 1.10.4,
$\left(\mathbf{A}^{\mathrm{T}}\right)_{ij} = \mathbf{e}_i \cdot \mathbf{A}^{\mathrm{T}}\mathbf{e}_j = \mathbf{e}_j \cdot \mathbf{Ae}_i = (\mathbf{A})_{ji} = A_{ji}$.

1.10.3 The Trace of a Tensor

The **trace** of a second order tensor **A**, denoted by $\mathrm{tr}\mathbf{A}$, is a scalar equal to the sum of the diagonal elements of its matrix representation. Thus (see Eqn. 1.4.3)

$$\boxed{\mathrm{tr}\mathbf{A} = A_{ii}} \qquad \textbf{Trace} \tag{1.10.5}$$

A more formal definition, again not relying on any particular coordinate system, is

$$\boxed{\mathrm{tr}\mathbf{A} = \mathbf{I} : \mathbf{A}} \qquad \textbf{Trace} \tag{1.10.6}$$

[1] as mentioned in §1.9, from the linearity of tensors, $\mathbf{uA} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{Av}$ and, for this reason, this expression is usually written simply as $\mathbf{uAv}$

and Eqn. 1.10.5 follows from 1.10.6 {▲Problem 4}. For the dyad $\mathbf{u} \otimes \mathbf{v}$ {▲Problem 5},

$$\mathrm{tr}(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \tag{1.10.7}$$

Another example is

$$\begin{aligned} \mathrm{tr}(\mathbf{E}^2) &= \mathbf{I} : \mathbf{E}^2 \\ &= \delta_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j) : E_{pq}E_{qr}(\mathbf{e}_p \otimes \mathbf{e}_r) \\ &= E_{iq}E_{qi} \end{aligned} \tag{1.10.8}$$

This and other important traces, and functions of the trace are listed here {▲Problem 6}:

$$\begin{aligned} \mathrm{tr}\mathbf{A} &= A_{ii} \\ \mathrm{tr}\mathbf{A}^2 &= A_{ij}A_{ji} \\ \mathrm{tr}\mathbf{A}^3 &= A_{ij}A_{jk}A_{ki} \\ (\mathrm{tr}\mathbf{A})^2 &= A_{ii}A_{jj} \\ (\mathrm{tr}\mathbf{A})^3 &= A_{ii}A_{jj}A_{kk} \end{aligned} \tag{1.10.9}$$

Some useful properties and relations involving the trace are {▲Problem 7}:

$$\begin{aligned} \mathrm{tr}\mathbf{A}^{\mathrm{T}} &= \mathrm{tr}\mathbf{A} \\ \mathrm{tr}(\mathbf{AB}) &= \mathrm{tr}(\mathbf{BA}) \\ \mathrm{tr}(\mathbf{A}+\mathbf{B}) &= \mathrm{tr}\mathbf{A} + \mathrm{tr}\mathbf{B} \\ \mathrm{tr}(\alpha\mathbf{A}) &= \alpha\mathrm{tr}\mathbf{A} \\ \mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) &= \mathrm{tr}(\mathbf{AB}^{\mathrm{T}}) = \mathrm{tr}(\mathbf{B}^{\mathrm{T}}\mathbf{A}) = \mathrm{tr}(\mathbf{BA}^{\mathrm{T}}) \end{aligned} \tag{1.10.10}$$

The double contraction of two tensors was earlier defined with respect to Cartesian coordinates, Eqn. 1.9.16. This last expression allows one to re-define the double contraction in terms of the trace, independent of any coordinate system.

Consider again the real vector space of second order tensors V^2 introduced in §1.8.5. The double contraction of two tensors as defined by 1.10.10e clearly satisfies the requirements of an inner product listed in §1.2.2. Thus this scalar quantity serves as an inner product for the space V^2:

$$\langle \mathbf{A}, \mathbf{B} \rangle \equiv \mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) \tag{1.10.11}$$

and generates an inner product space.

Just as the base vectors $\{\mathbf{e}_i\}$ form an orthonormal set in the inner product (vector dot product) of the space of vectors V, so the base dyads $\{\mathbf{e}_i \otimes \mathbf{e}_j\}$ form an orthonormal set in the inner product 1.10.11 of the space of second order tensors V^2. For example,

$$\langle \mathbf{e}_1 \otimes \mathbf{e}_1, \mathbf{e}_1 \otimes \mathbf{e}_1 \rangle = (\mathbf{e}_1 \otimes \mathbf{e}_1) : (\mathbf{e}_1 \otimes \mathbf{e}_1) = 1 \tag{1.10.12}$$

Similarly, just as the dot product is zero for orthogonal vectors, when the double contraction of two tensors **A** and **B** is zero, one says that the tensors are **orthogonal**,

$$\mathbf{A} : \mathbf{B} = \mathrm{tr}\left(\mathbf{A}^{\mathrm{T}}\mathbf{B}\right) = 0, \qquad \mathbf{A}, \mathbf{B} \text{ orthogonal} \tag{1.10.13}$$

1.10.4 The Norm of a Tensor

Using 1.2.8 and 1.10.11, the **norm** of a second order tensor **A**, denoted by $|\mathbf{A}|$ (or $\|\mathbf{A}\|$), is defined by

$$|\mathbf{A}| = \sqrt{\mathbf{A} : \mathbf{A}} \tag{1.10.14}$$

This is analogous to the norm $|\mathbf{a}|$ of a vector **a**, $\sqrt{\mathbf{a} \cdot \mathbf{a}}$.

1.10.5 The Determinant of a Tensor

The **determinant** of a second order tensor **A** is defined to be the determinant of the matrix $[\mathbf{A}]$ of components of the tensor:

$$\begin{aligned} \det \mathbf{A} &= \det \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \\ &= \varepsilon_{ijk} A_{i1} A_{j2} A_{k3} \\ &= \varepsilon_{ijk} A_{1i} A_{2j} A_{3k} \end{aligned} \tag{1.10.15}$$

Some useful properties of the determinant are {▲Problem 8}

$$\begin{aligned} \det(\mathbf{AB}) &= \det \mathbf{A} \det \mathbf{B} \\ \det \mathbf{A}^{\mathrm{T}} &= \det \mathbf{A} \\ \det(\alpha \mathbf{A}) &= \alpha^3 \det \mathbf{A} \\ \det(\mathbf{u} \otimes \mathbf{v}) &= 0 \\ \varepsilon_{pqr}(\det \mathbf{A}) &= \varepsilon_{ijk} A_{ip} A_{jq} A_{kr} \\ (\mathbf{Ta} \times \mathbf{Tb}) \cdot \mathbf{Tc} &= (\det \mathbf{T})[(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}] \end{aligned} \tag{1.10.16}$$

Note that $\det \mathbf{A}$, like $\mathrm{tr}\mathbf{A}$, is independent of the choice of coordinate system / basis.

1.10.6 The Inverse of a Tensor

The **inverse** of a second order tensor **A**, denoted by $\mathbf{A}^{-1}$, is defined by

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A} \tag{1.10.17}$$

The inverse of a tensor exists only if it is **non-singular** (a **singular** tensor is one for which $\det \mathbf{A} = 0$), in which case it is said to be **invertible**.

Some useful properties and relations involving the inverse are:

$$\begin{aligned} (\mathbf{A}^{-1})^{-1} &= \mathbf{A} \\ (\alpha\mathbf{A})^{-1} &= (1/\alpha)\mathbf{A}^{-1} \\ (\mathbf{A}\mathbf{B})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} \\ \det(\mathbf{A}^{-1}) &= (\det \mathbf{A})^{-1} \end{aligned} \tag{1.10.18}$$

Since the inverse of the transpose is equivalent to the transpose of the inverse, the following notation is used:

$$\mathbf{A}^{-\mathrm{T}} \equiv (\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1} \tag{1.10.19}$$

1.10.7 Orthogonal Tensors

An **orthogonal** tensor **Q** is a linear vector transformation satisfying the condition

$$\mathbf{Qu} \cdot \mathbf{Qv} = \mathbf{u} \cdot \mathbf{v} \tag{1.10.20}$$

for all vectors **u** and **v**. Thus **u** is transformed to $\mathbf{Qu}$, **v** is transformed to $\mathbf{Qv}$ and the dot product $\mathbf{u} \cdot \mathbf{v}$ is invariant under the transformation. Thus the magnitude of the vectors and the angle between the vectors is preserved, Fig. 1.10.1.

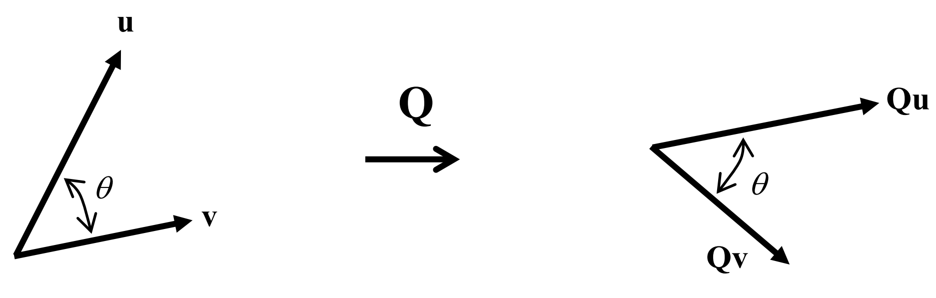


Figure 1.10.1: An orthogonal tensor

Since

$$\mathbf{Qu} \cdot \mathbf{Qv} = \mathbf{u}\mathbf{Q}^{\mathrm{T}} \cdot \mathbf{Qv} = \mathbf{u} \cdot \left(\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\right) \cdot \mathbf{v} \tag{1.10.21}$$

it follows that for $\mathbf{u} \cdot \mathbf{v}$ to be preserved under the transformation, $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$, which is also used as the definition of an orthogonal tensor. Some useful properties of orthogonal tensors are {▲Problem 10}:

$$\begin{aligned} \mathbf{Q}\mathbf{Q}^{\mathrm{T}} &= \mathbf{I} = \mathbf{Q}^{\mathrm{T}}\mathbf{Q}, \qquad Q_{ik}Q_{jk} = \delta_{ij} = Q_{ki}Q_{kj} \\ \mathbf{Q}^{-1} &= \mathbf{Q}^{\mathrm{T}} \\ \det\mathbf{Q} &= \pm 1 \end{aligned} \tag{1.10.22}$$

1.10.8 Rotation Tensors

If for an orthogonal tensor, $\det\mathbf{Q} = +1$, $\mathbf{Q}$ is said to be a **proper** orthogonal tensor, corresponding to a **rotation**. If $\det\mathbf{Q} = -1$, $\mathbf{Q}$ is said to be an **improper** orthogonal tensor, corresponding to a **reflection**. Proper orthogonal tensors are also called **rotation tensors**.

1.10.9 Change of Basis Tensors

Consider a rotation tensor $\mathbf{Q}$ which rotates the base vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ into a second set, $\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'$, Fig. 1.10.2.

$$\mathbf{e}_i' = \mathbf{Q}\mathbf{e}_i \qquad i = 1, 2, 3 \tag{1.10.23}$$

Such a tensor can be termed a **change of basis tensor** from $\{\mathbf{e}_i\}$ to $\{\mathbf{e}_i'\}$. The transpose $\mathbf{Q}^{\mathrm{T}}$ rotates the base vectors $\mathbf{e}_i'$ back to $\mathbf{e}_i$ and is thus **change of basis tensor** from $\{\mathbf{e}_i'\}$ to $\{\mathbf{e}_i\}$. The components of $\mathbf{Q}$ in the $\mathbf{e}_i$ coordinate system are, from 1.9.4, $Q_{ij} = \mathbf{e}_i\mathbf{Q}\mathbf{e}_j$ and so, from 1.10.23,

$$\mathbf{Q} = Q_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \qquad Q_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j', \tag{1.10.24}$$

which are the direction cosines between the axes (see Fig. 1.5.4).

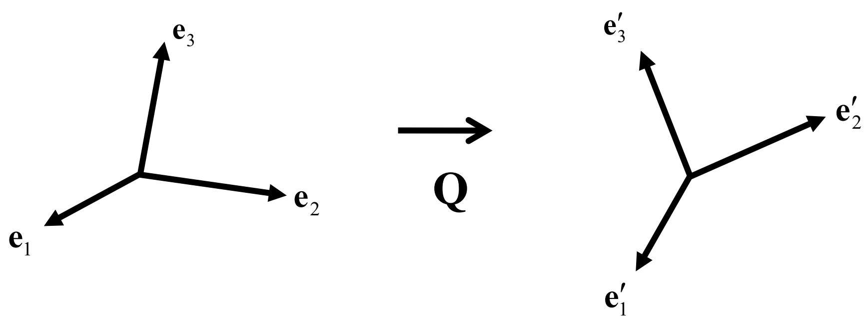


Figure 1.10.2: Rotation of a set of base vectors

The change of basis tensor can also be expressed in terms of the base vectors from *both* bases:

$$\mathbf{Q} = \mathbf{e}'_i \otimes \mathbf{e}_i, \tag{1.10.25}$$

from which the above relations can easily be derived, for example $\mathbf{e}'_i = \mathbf{Q}\mathbf{e}_i$, $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$, etc.

Consider now the operation of the change of basis tensor on a vector:

$$\mathbf{Q}\mathbf{v} = v_i(\mathbf{Q}\mathbf{e}_i) = v_i\mathbf{e}'_i \tag{1.10.26}$$

Thus $\mathbf{Q}$ transforms $\mathbf{v}$ into a second vector $\mathbf{v}'$, but this new vector has the *same components* with respect to the basis $\mathbf{e}'_i$, as $\mathbf{v}$ has with respect to the basis $\mathbf{e}_i$, $v'_i = v_i$. Note the difference between this and the coordinate transformations of §1.5: here there are two different vectors, $\mathbf{v}$ and $\mathbf{v}'$.

Example

Consider the two-dimensional rotation tensor

$$\mathbf{Q} = \begin{bmatrix} 0 & -1 \\ +1 & 0 \end{bmatrix}(\mathbf{e}_i \otimes \mathbf{e}_j) \equiv \mathbf{e}'_i \otimes \mathbf{e}_i$$

which corresponds to a rotation of the base vectors through $\pi/2$. The vector $\mathbf{v} = [1 \quad 1]^{\mathrm{T}}$ then transforms into (see Fig. 1.10.3)

$$\mathbf{Q}\mathbf{v} = \begin{bmatrix} -1 \\ +1 \end{bmatrix}\mathbf{e}_i = \begin{bmatrix} +1 \\ +1 \end{bmatrix}\mathbf{e}'_i$$

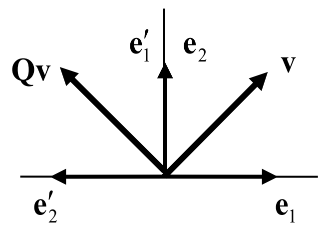


Figure 1.10.3: a rotated vector

■

Similarly, for a second order tensor $\mathbf{A}$, the operation

$$\mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}(A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j)\mathbf{Q}^{\mathrm{T}} = A_{ij}(\mathbf{Q}\mathbf{e}_i \otimes \mathbf{e}_j\mathbf{Q}^{\mathrm{T}}) = A_{ij}(\mathbf{Q}\mathbf{e}_i \otimes \mathbf{Q}\mathbf{e}_j) = A_{ij}\mathbf{e}'_i \otimes \mathbf{e}'_j \tag{1.10.27}$$

results in a new tensor which has the same components with respect to the $\mathbf{e}'_i$, as $\mathbf{A}$ has with respect to the $\mathbf{e}_i$, $A'_{ij} = A_{ij}$.

1.10.10 Symmetric and Skew Tensors

A tensor $\mathbf{T}$ is said to be **symmetric** if it is identical to the transposed tensor, $\mathbf{T} = \mathbf{T}^{\mathrm{T}}$, and **skew** (**antisymmetric**) if $\mathbf{T} = -\mathbf{T}^{\mathrm{T}}$.

Any tensor $\mathbf{A}$ can be (uniquely) decomposed into a symmetric tensor $\mathbf{S}$ and a skew tensor $\mathbf{W}$, where

$$\begin{aligned} \mathrm{sym}\mathbf{A} &\equiv \mathbf{S} = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right) \\ \mathrm{skew}\mathbf{A} &\equiv \mathbf{W} = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right) \end{aligned} \tag{1.10.28}$$

and

$$\mathbf{S} = \mathbf{S}^{\mathrm{T}}, \qquad \mathbf{W} = -\mathbf{W}^{\mathrm{T}} \tag{1.10.29}$$

In matrix notation one has

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{bmatrix}, \qquad [W] = \begin{bmatrix} 0 & W_{12} & W_{13} \\ -W_{12} & 0 & W_{23} \\ -W_{13} & -W_{23} & 0 \end{bmatrix} \tag{1.10.30}$$

Some useful properties of symmetric and skew tensors are {▲Problem 13}:

$$\begin{aligned} &\mathbf{S} : \mathbf{B} = \mathbf{S} : \mathbf{B}^{\mathrm{T}} = \mathbf{S} : \tfrac{1}{2}\left(\mathbf{B} + \mathbf{B}^{\mathrm{T}}\right) \\ &\mathbf{W} : \mathbf{B} = -\mathbf{W} : \mathbf{B}^{\mathrm{T}} = \mathbf{W} : \tfrac{1}{2}\left(-\mathbf{B}^{\mathrm{T}}\right) \\ &\mathbf{S} : \mathbf{W} = 0 \\ &\mathrm{tr}(\mathbf{SW}) = 0 \\ &\mathbf{v} \cdot \mathbf{W}\mathbf{v} = 0 \\ &\det \mathbf{W} = 0 \quad (\text{has no inverse}) \end{aligned} \tag{1.10.31}$$

where $\mathbf{v}$ and $\mathbf{B}$ denote any arbitrary vector and second-order tensor respectively.

Note that symmetry and skew-symmetry are tensor properties, independent of coordinate system.

1.10.11 Axial Vectors

A skew tensor **W** has only three independent coefficients, so it behaves "like a vector" with three components. Indeed, a skew tensor can always be written in the form

$$\mathbf{Wu} = \boldsymbol{\omega} \times \mathbf{u} \tag{1.10.32}$$

where **u** is any vector and $\boldsymbol{\omega}$ characterises the **axial** (or **dual**) vector of the skew tensor **W**. The components of **W** can be obtained from the components of $\boldsymbol{\omega}$ through

$$\begin{aligned} W_{ij} &= \mathbf{e}_i \cdot \mathbf{W}\mathbf{e}_j = \mathbf{e}_i \cdot (\boldsymbol{\omega} \times \mathbf{e}_j) = \mathbf{e}_i \cdot (\omega_k \mathbf{e}_k \times \mathbf{e}_j) \\ &= \mathbf{e}_i \cdot (\omega_k \varepsilon_{kjp} \mathbf{e}_p) = \varepsilon_{kji} \omega_k \\ &= -\varepsilon_{ijk} \omega_k \end{aligned} \tag{1.10.33}$$

If one knows the components of **W**, one can find the components of $\boldsymbol{\omega}$ by inverting this equation, whence {▲Problem 14}

$$\boldsymbol{\omega} = -W_{23}\mathbf{e}_1 + W_{13}\mathbf{e}_2 - W_{12}\mathbf{e}_3 \tag{1.10.34}$$

Example (of an Axial Vector)

Decompose the tensor

$$\mathbf{T} = [T_{ij}] = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

into its symmetric and skew parts. Also find the axial vector for the skew part. Verify that $\mathbf{Wa} = \boldsymbol{\omega} \times \mathbf{a}$ for $\mathbf{a} = \mathbf{e}_1 + \mathbf{e}_3$.

<u>Solution</u>

One has

$$\mathbf{S} = \frac{1}{2}[\mathbf{T} + \mathbf{T}^{\mathrm{T}}] = \frac{1}{2}\left\{ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 4 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} \right\} = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

$$\mathbf{W} = \frac{1}{2}[\mathbf{T} - \mathbf{T}^{\mathrm{T}}] = \frac{1}{2}\left\{ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 4 & 1 \\ 2 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} \right\} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

The axial vector is

$$\boldsymbol{\omega} = -W_{23}\mathbf{e}_1 + W_{13}\mathbf{e}_2 - W_{12}\mathbf{e}_3 = \mathbf{e}_2 + \mathbf{e}_3$$

and it can be seen that

$$\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{or} \quad \begin{aligned} \mathbf{Wa} &= W_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)(\mathbf{e}_1 + \mathbf{e}_3) = W_{ij}(\delta_{j1} + \delta_{j3})\mathbf{e}_i = (W_{i1} + W_{i3})\mathbf{e}_i \\ &= (W_{11} + W_{13})\mathbf{e}_1 + (W_{21} + W_{23})\mathbf{e}_2 + (W_{31} + W_{33})\mathbf{e}_3 \\ &= \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 \end{aligned}$$

and

$$\boldsymbol{\omega} \times \mathbf{a} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} = \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$$

■

The Spin Tensor

The velocity of a particle rotating in a rigid body motion is given by $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{x}$, where $\boldsymbol{\omega}$ is the angular velocity vector and **x** is the position vector relative to the origin on the axis of rotation (see Problem 9, §1.1). If the velocity can be written in terms of a skew-symmetric second order tensor $\mathbf{w}$, such that $\mathbf{wx} = \mathbf{v}$, then it follows from $\mathbf{wx} = \boldsymbol{\omega} \times \mathbf{x}$ that the angular velocity vector $\boldsymbol{\omega}$ is the axial vector of $\mathbf{w}$. In this context, $\mathbf{w}$ is called the **spin tensor**.

1.10.12 Spherical and Deviatoric Tensors

Every tensor **A** can be decomposed into its so-called **spherical** part and its **deviatoric** part, i.e.

$$\mathbf{A} = \text{sph}\mathbf{A} + \text{dev}\mathbf{A} \tag{1.10.35}$$

where

$$\begin{aligned} \text{sph}\mathbf{A} &= \tfrac{1}{3}(\text{tr}\mathbf{A})\mathbf{I} \\ &= \begin{bmatrix} \tfrac{1}{3}(A_{11} + A_{22} + A_{33}) & 0 & 0 \\ 0 & \tfrac{1}{3}(A_{11} + A_{22} + A_{33}) & 0 \\ 0 & 0 & \tfrac{1}{3}(A_{11} + A_{22} + A_{33}) \end{bmatrix} \\ \text{dev}\mathbf{A} &= \mathbf{A} - \text{sph}\mathbf{A} \\ &= \begin{bmatrix} A_{11} - \tfrac{1}{3}(A_{11} + A_{22} + A_{33}) & A_{12} & A_{13} \\ A_{21} & A_{22} - \tfrac{1}{3}(A_{11} + A_{22} + A_{33}) & A_{23} \\ A_{31} & A_{32} & A_{33} - \tfrac{1}{3}(A_{11} + A_{22} + A_{33}) \end{bmatrix} \end{aligned} \tag{1.10.36}$$

Any tensor of the form $\alpha\mathbf{I}$ is known as a **spherical tensor**, while $\text{dev}\mathbf{A}$ is known as a deviator of **A**, or a **deviatoric tensor**.

Some important properties of the spherical and deviatoric tensors are

$$\begin{aligned} \text{tr}(\text{dev}\mathbf{A}) &= 0 \\ \text{sph}(\text{dev}\mathbf{A}) &= 0 \\ \text{dev}\mathbf{A} : \text{sph}\mathbf{B} &= 0 \end{aligned} \qquad (1.10.37)$$

1.10.13 Positive Definite Tensors

A **positive definite** tensor **A** is one which satisfies the relation

$$\mathbf{vAv} > 0, \qquad \forall\, \mathbf{v} \neq \mathbf{o} \qquad (1.10.38)$$

The tensor is called **positive semi-definite** if $\mathbf{vAv} \geq 0$.

In component form,

$$v_i A_{ij} v_j = A_{11}v_1^2 + A_{12}v_1v_2 + A_{13}v_1v_3 + A_{21}v_2v_1 + A_{22}v_2^2 + \cdots \qquad (1.10.39)$$

and so the diagonal elements of the matrix representation of a positive definite tensor must always be positive.

It can be shown that the following conditions are necessary for a tensor **A** to be positive definite (although they are not sufficient):

(i) the diagonal elements of $[\mathbf{A}]$ are positive
(ii) the largest element of $[\mathbf{A}]$ lies along the diagonal
(iii) $\det\mathbf{A} > 0$
(iv) $A_{ii} + A_{jj} > 2A_{ij}$ for $i \neq j$ (no sum over i, j)

These conditions are seen to hold for the following matrix representation of an example positive definite tensor:

$$[\mathbf{A}] = \begin{bmatrix} 2 & 2 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A necessary and sufficient condition for a tensor to be positive definite is given in the next section, during the discussion of the eigenvalue problem.

One of the key properties of a positive definite tensor is that, since $\det\mathbf{A} > 0$, positive definite tensors are always invertible.

An alternative definition of positive definiteness is the equivalent expression

$$\mathbf{A} : \mathbf{v} \otimes \mathbf{v} > 0 \tag{1.10.40}$$

1.10.14 Problems

1. Show that the components of the (second-order) identity tensor are given by $I_{ij} = \delta_{ij}$.
2. Show that
   (a) $(\mathbf{u} \otimes \mathbf{v})\mathbf{A} = \mathbf{u} \otimes (\mathbf{A}^{\mathrm{T}}\mathbf{v})$
   (b) $\mathbf{A} : (\mathbf{BC}) = (\mathbf{B}^{\mathrm{T}}\mathbf{A}) : \mathbf{C} = (\mathbf{AC}^{\mathrm{T}}) : \mathbf{B}$
3. Use (1.10.4) to show that $\mathbf{I}^{\mathrm{T}} = \mathbf{I}$.
4. Show that (1.10.6) implies (1.10.5) for the trace of a tensor.
5. Show that $\mathrm{tr}(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$.
6. Formally derive the index notation for the functions
   $\mathrm{tr}\mathbf{A}^2, \quad \mathrm{tr}\mathbf{A}^3, \quad (\mathrm{tr}\mathbf{A})^2, \quad (\mathrm{tr}\mathbf{A})^3$
7. Show that $\mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B})$.
8. Prove (1.10.16f), $(\mathbf{Ta} \times \mathbf{Tb}) \cdot \mathbf{Tc} = (\det \mathbf{T})[(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]$.
9. Show that $(\mathbf{A}^{-1})^{\mathrm{T}} : \mathbf{A} = 3$. [Hint: one way of doing this is using the result from Problem 7.]
10. Use 1.10.16b and 1.10.18d to prove 1.10.22c, $\det \mathbf{Q} = \pm 1$.
11. Use the explicit dyadic representation of the rotation tensor, $\mathbf{Q} = \mathbf{e}'_i \otimes \mathbf{e}_i$, to show that the components of $\mathbf{Q}$ in the "second", $ox'_1x'_2x'_3$, coordinate system are the same as those in the first system [hint: use the rule $Q'_{ij} = \mathbf{e}'_i \cdot \mathbf{Q}\mathbf{e}'_j$]
12. Consider the tensor $\mathbf{D}$ with components (in a certain coordinate system)
   $$\begin{bmatrix} 1/\sqrt{2} & 1/2 & -1/2 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/2 & 1/2 \end{bmatrix}$$
   Show that $\mathbf{D}$ is a rotation tensor (just show that $\mathbf{D}$ is proper orthogonal).
13. Show that $\mathrm{tr}(\mathbf{SW}) = 0$.
14. Multiply across (1.10.32), $W_{ij} = -\varepsilon_{ijk}\omega_k$, by ε_{ijp} to show that $\boldsymbol{\omega} = -\frac{1}{2}\varepsilon_{ijk}W_{ij}\mathbf{e}_k$. [Hint: use the relation 1.3.19b, $\varepsilon_{ijp}\varepsilon_{ijk} = 2\delta_{pk}$.]
15. Show that $\frac{1}{2}(\mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a})$ is a skew tensor $\mathbf{W}$. Show that its axial vector is $\boldsymbol{\omega} = \frac{1}{2}(\mathbf{b} \times \mathbf{a})$. [Hint: first prove that $(\mathbf{b} \cdot \mathbf{u})\mathbf{a} - (\mathbf{a} \cdot \mathbf{u})\mathbf{b} = \mathbf{u} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{b} \times \mathbf{a}) \times \mathbf{u}$.]
16. Find the spherical and deviatoric parts of the tensor $\mathbf{A}$ for which $A_{ij} = 1$.

1.11 The Eigenvalue Problem and Polar Decomposition

1.11.1 Eigenvalues, Eigenvectors and Invariants of a Tensor

Consider a second-order tensor **A**. Suppose that one can find a scalar λ and a (non-zero) normalised, i.e. unit, vector $\hat{\mathbf{n}}$ such that

$$\mathbf{A}\hat{\mathbf{n}} = \lambda\hat{\mathbf{n}} \tag{1.11.1}$$

In other words, **A** transforms the vector $\hat{\mathbf{n}}$ into a vector parallel to itself, Fig. 1.11.1. If this transformation is possible, the scalars are called the **eigenvalues** (or **principal values**) of the tensor, and the vectors are called the **eigenvectors** (or **principal directions** or **principal axes**) of the tensor. It will be seen that there are *three* vectors $\hat{\mathbf{n}}$ (to each of which corresponds some scalar λ) for which the above holds.

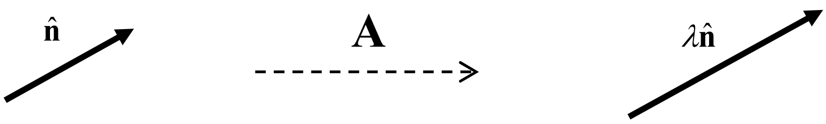


Figure 1.11.1: the action of a tensor A on a unit vector

Equation 1.11.1 can be solved for the eigenvalues and eigenvectors by rewriting it as

$$(\mathbf{A} - \lambda\mathbf{I})\hat{\mathbf{n}} = 0 \tag{1.11.2}$$

or, in terms of a Cartesian coordinate system,

$$\begin{aligned}
&A_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)\hat{n}_k\mathbf{e}_k - \lambda\delta_{pq}(\mathbf{e}_p \otimes \mathbf{e}_q)\hat{n}_r\mathbf{e}_r = 0 \\
&\rightarrow A_{ij}\hat{n}_j\mathbf{e}_i - \lambda\hat{n}_r\mathbf{e}_r = 0 \\
&\rightarrow (A_{ij}\hat{n}_j - \lambda\hat{n}_i)\mathbf{e}_i = 0
\end{aligned}$$

In full,

$$\begin{aligned}
&[(A_{11} - \lambda)\hat{n}_1 + A_{12}\hat{n}_2 + A_{13}\hat{n}_3]\mathbf{e}_1 = 0 \\
&[A_{21}\hat{n}_1 + (A_{22} - \lambda)\hat{n}_2 + A_{23}\hat{n}_3]\mathbf{e}_2 = 0 \\
&[A_{31}\hat{n}_1 + A_{32}\hat{n}_2 + (A_{33} - \lambda)\hat{n}_3]\mathbf{e}_3 = 0
\end{aligned} \tag{1.11.3}$$

Dividing out the base vectors, this is a set of three homogeneous equations in three unknowns (if one treats λ as known). From basic linear algebra, this system has a solution (apart from $\hat{n}_i = 0$) if and only if the determinant of the coefficient matrix is zero, i.e. if

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \det\begin{bmatrix} A_{11}-\lambda & A_{12} & A_{13} \\ A_{21} & A_{22}-\lambda & A_{23} \\ A_{31} & A_{32} & A_{33}-\lambda \end{bmatrix} = 0 \tag{1.11.4}$$

Evaluating the determinant, one has the following cubic **characteristic equation** of **A**,

$$\boxed{\lambda^3 - \mathrm{I}_\mathbf{A}\lambda^2 + \mathrm{II}_\mathbf{A}\lambda - \mathrm{III}_\mathbf{A} = 0} \quad \textbf{Tensor Characteristic Equation} \tag{1.11.5}$$

where

$$\begin{aligned} \mathrm{I}_\mathbf{A} &= A_{ii} \\ &= \mathrm{tr}\mathbf{A} \\ \mathrm{II}_\mathbf{A} &= \tfrac{1}{2}\left(A_{ii}A_{jj} - A_{ji}A_{ij}\right) \\ &= \tfrac{1}{2}\left[(\mathrm{tr}\mathbf{A})^2 - \mathrm{tr}(\mathbf{A}^2)\right] \\ \mathrm{III}_\mathbf{A} &= \varepsilon_{ijk}A_{1i}A_{2j}A_{3k} \\ &= \det\mathbf{A} \end{aligned} \tag{1.11.6}$$

It can be seen that there are three roots $\lambda_1, \lambda_2, \lambda_3$, to the characteristic equation. Solving for λ, one finds that

$$\begin{aligned} \mathrm{I}_\mathbf{A} &= \lambda_1 + \lambda_2 + \lambda_3 \\ \mathrm{II}_\mathbf{A} &= \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 \\ \mathrm{III}_\mathbf{A} &= \lambda_1\lambda_2\lambda_3 \end{aligned} \tag{1.11.7}$$

The eigenvalues (principal values) λ_i must be independent of any coordinate system and, from Eqn. 1.11.5, it follows that the functions $\mathrm{I}_\mathbf{A}, \mathrm{II}_\mathbf{A}, \mathrm{III}_\mathbf{A}$ are also independent of any coordinate system. They are called the **principal scalar invariants** (or simply **invariants**) of the tensor.

Once the eigenvalues are found, the eigenvectors (principal directions) can be found by solving

$$\begin{aligned} (A_{11}-\lambda)\hat{n}_1 + A_{12}\hat{n}_2 + A_{13}\hat{n}_3 &= 0 \\ A_{21}\hat{n}_1 + (A_{22}-\lambda)\hat{n}_2 + A_{23}\hat{n}_3 &= 0 \\ A_{31}\hat{n}_1 + A_{32}\hat{n}_2 + (A_{33}-\lambda)\hat{n}_3 &= 0 \end{aligned} \tag{1.11.8}$$

for the three components of the principal direction vector $\hat{n}_1, \hat{n}_2, \hat{n}_3$, in addition to the condition that $\hat{\mathbf{n}}\cdot\hat{\mathbf{n}} = \hat{n}_i\hat{n}_i = 1$. There will be three vectors $\hat{\mathbf{n}} = \hat{n}_i\mathbf{e}_i$, one corresponding to each of the three principal values.

Note: a unit eigenvector $\hat{\mathbf{n}}$ has been used in the above discussion, but *any* vector parallel to $\hat{\mathbf{n}}$, for example $\alpha\hat{\mathbf{n}}$, is also an eigenvector (with the same eigenvalue λ):

$$\mathbf{A}(\alpha\hat{\mathbf{n}}) = \alpha(\mathbf{A}\hat{\mathbf{n}}) = \alpha(\lambda\hat{\mathbf{n}}) = \lambda(\alpha\hat{\mathbf{n}})$$

Example (of Eigenvalues and Eigenvectors of a Tensor)

A second order tensor **T** is given with respect to the axes $Ox_1x_2x_3$ by the values

$$\mathbf{T} = [\mathbf{T}]_{ij} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -6 & -12 \\ 0 & -12 & 1 \end{bmatrix}.$$

Determine (a) the principal values, (b) the principal directions (and sketch them).

<u>Solution:</u>
(a)
The principal values are the solution to the characteristic equation

$$\begin{vmatrix} 5-\lambda & 0 & 0 \\ 0 & -6-\lambda & -12 \\ 0 & -12 & 1-\lambda \end{vmatrix} = (-10+\lambda)(5-\lambda)(15+\lambda) = 0$$

which yields the three principal values $\lambda_1 = 10,\ \lambda_2 = 5,\ \lambda_3 = -15$.
(b)
The eigenvectors are now obtained from $(T_{ij} - \delta_{ij}\lambda)n_j = 0$. First, for $\lambda_1 = 10$,

$$\begin{aligned} -5n_1 + 0n_2 + 0n_3 &= 0 \\ 0n_1 - 16n_2 - 12n_3 &= 0 \\ 0n_1 - 12n_2 - 9n_3 &= 0 \end{aligned}$$

and using also the equation $n_i n_i = 1$ leads to $\hat{\mathbf{n}}_1 = -(3/5)\mathbf{e}_2 + (4/5)\mathbf{e}_3$. Similarly, for $\lambda_2 = 5$ and $\lambda_3 = -15$, one has, respectively,

$$\begin{aligned} 0n_1 + 0n_2 + 0n_3 &= 0 \\ 0n_1 - 11n_2 - 12n_3 &= 0 \\ 0n_1 - 12n_2 - 4n_3 &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} 20n_1 + 0n_2 + 0n_3 &= 0 \\ 0n_1 + 9n_2 - 12n_3 &= 0 \\ 0n_1 - 12n_2 + 16n_3 &= 0 \end{aligned}$$

which yield $\hat{\mathbf{n}}_2 = \mathbf{e}_1$ and $\hat{\mathbf{n}}_3 = (4/5)\mathbf{e}_2 + (3/5)\mathbf{e}_3$. The principal directions are sketched in Fig. 1.11.2.

Note: the three components of a principal direction, n_1, n_2, n_3, are the direction cosines between that direction and the three coordinate axes respectively. For example, for λ_1 with $n_1 = 0, n_2 = -3/5, n_3 = 4/5$, the angles made with the coordinate axes x_1, x_2, x_3, are 0, 127° and 37°.

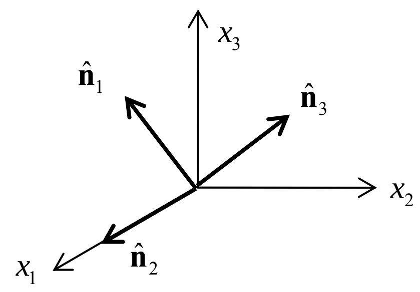


Figure 1.11.2: eigenvectors of the tensor T

■

1.11.2 Real Symmetric Tensors

Suppose now that **A** is a *real symmetric* tensor (real meaning that its components are real). In that case it can be proved (see below) that[1]

(i) the eigenvalues are real
(ii) the three eigenvectors form an orthonormal basis $\{\hat{\mathbf{n}}_i\}$.

In that case, the components of **A** can be written relative to the basis of principal directions as (see Fig. 1.11.3)

$$\mathbf{A} = A_{ij}\left(\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_j\right) \tag{1.11.9}$$

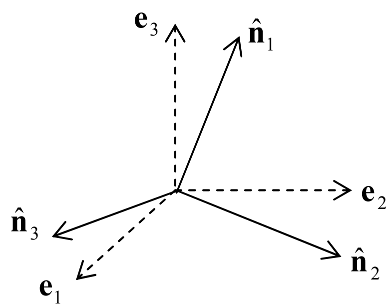


Figure 1.11.3: eigenvectors forming an orthonormal set

The components of **A** in this new basis can be obtained from Eqn. 1.9.4,

$$\begin{aligned} A_{ij} &= \hat{\mathbf{n}}_i \cdot \mathbf{A}\hat{\mathbf{n}}_j \\ &= \hat{\mathbf{n}}_i \cdot \left(\lambda_j \hat{\mathbf{n}}_j\right) \qquad \text{(no summation over } j\text{)} \\ &= \begin{cases} \lambda_i, & i = j \\ 0, & i \neq j \end{cases} \end{aligned} \tag{1.11.10}$$

where λ_i is the eigenvalue corresponding to the basis vector $\hat{\mathbf{n}}_i$. Thus[2]

[1] this was the case in the previous example – the tensor is real symmetric and the principal directions are orthogonal

$$\boxed{\mathbf{A} = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i} \qquad \textbf{Spectral Decomposition} \qquad (1.11.11)$$

This is called the **spectral decomposition** (or **spectral representation**) of **A**. In matrix form,

$$[\mathbf{A}] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \qquad (1.11.12)$$

For example, the tensor used in the previous example can be written in terms of the basis vectors in the principal directions as

$$\mathbf{T} = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -15 \end{bmatrix}, \qquad \text{basis: } \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_j$$

To prove that real symmetric tensors have real eigenvalues and orthonormal eigenvectors, take $\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2, \hat{\mathbf{n}}_3$ to be the eigenvectors of an arbitrary tensor **A**, with components $\hat{n}_{1i}, \hat{n}_{2i}, \hat{n}_{3i}$, which are solutions of (the 9 equations – see Eqn. 1.11.2)

$$\begin{aligned} (\mathbf{A} - \lambda_1 \mathbf{I})\hat{\mathbf{n}}_1 &= 0 \\ (\mathbf{A} - \lambda_2 \mathbf{I})\hat{\mathbf{n}}_2 &= 0 \\ (\mathbf{A} - \lambda_3 \mathbf{I})\hat{\mathbf{n}}_3 &= 0 \end{aligned} \qquad (1.11.13)$$

Dotting the first of these by $\hat{\mathbf{n}}_1$ and the second by $\hat{\mathbf{n}}_1$, leads to

$$\begin{aligned} (\mathbf{A}\mathbf{n}_1) \cdot \mathbf{n}_2 - \lambda_1 \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 &= 0 \\ (\mathbf{A}\mathbf{n}_2) \cdot \mathbf{n}_1 - \lambda_2 \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 &= 0 \end{aligned}$$

Using the fact that $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$, subtracting these equations leads to

$$(\lambda_2 - \lambda_1)\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 = 0 \qquad (1.11.14)$$

Assume now that the eigenvalues are not all real. Since the coefficients of the characteristic equation are all real, this implies that the eigenvalues come in a complex conjugate pair, say λ_1 and λ_2, and one real eigenvalue λ_3. It follows from Eqn. 1.11.13 that the components of $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are conjugates of each other, say $\hat{\mathbf{n}}_1 = \mathbf{a} + \mathbf{b}i$, $\hat{\mathbf{n}}_2 = \mathbf{a} - \mathbf{b}i$, and so

[2] it is necessary to introduce the summation sign here, because the summation convention is only used when *two* indices are the same – it cannot be used when there are more than two indices the same

$$\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 = (\mathbf{a} + \mathbf{b}i) \cdot (\mathbf{a} - \mathbf{b}i) = |\mathbf{a}|^2 + |\mathbf{b}|^2 > 0$$

It follows from 1.11.14 that $\lambda_2 - \lambda_1 = 0$ which is a contradiction, since this cannot be true for conjugate pairs. Thus the original assumption regarding complex roots must be false and the eigenvalues are all real. With three distinct eigenvalues, Eqn. 1.11.14 (and similar) show that the eigenvectors form an orthonormal set. When the eigenvalues are not distinct, more than one set of eigenvectors may be taken to form an orthonormal set (see the next subsection).

Equal Eigenvalues

There are some special tensors for which two or three of the principal directions are equal. When all three are equal, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, one has $\mathbf{A} = \lambda\mathbf{I}$, and the tensor is spherical: every direction is a principal direction, since $\mathbf{A}\hat{\mathbf{n}} = \lambda\mathbf{I}\hat{\mathbf{n}} = \lambda\hat{\mathbf{n}}$ for all $\hat{\mathbf{n}}$. When two of the eigenvalues are equal, one of the eigenvectors will be unique but the other two directions will be arbitrary – one can choose any two principal directions in the plane perpendicular to the uniquely determined direction, in order to form an orthonormal set.

Eigenvalues and Positive Definite Tensors

Since $\mathbf{A}\hat{\mathbf{n}} = \lambda\hat{\mathbf{n}}$, then $\hat{\mathbf{n}} \cdot \mathbf{A}\hat{\mathbf{n}} = \hat{\mathbf{n}} \cdot \lambda\hat{\mathbf{n}} = \lambda$. Thus if **A** is positive definite, Eqn. 1.10.38, the eigenvalues are all *positive*.

In fact, it can be shown that a tensor is positive definite if and only if its symmetric part has all positive eigenvalues.

Note: if there exists a non-zero eigenvector corresponding to a zero eigenvalue, then the tensor is singular. This is the case for the skew tensor **W**, which is singular. Since $\mathbf{W}\boldsymbol{\omega} = \boldsymbol{\omega} \times \boldsymbol{\omega} = \mathbf{o} = 0\boldsymbol{\omega}$ (see , §1.10.11), the axial vector $\boldsymbol{\omega}$ is an eigenvector corresponding to a zero eigenvalue of **W**.

1.11.3 Maximum and Minimum Values

The diagonal components of a tensor **A**, A_{11}, A_{22} A_{33}, have different values in different coordinate systems. However, the three eigenvalues include the extreme (maximum and minimum) possible values that any of these three components can take, in any coordinate system. To prove this, consider an arbitrary set of unit base vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, other than the eigenvectors. From Eqn. 1.9.4, the components of **A** in a new coordinate system with these base vectors are $A'_{ij} = \mathbf{e}_i \mathbf{A} \mathbf{e}_j$ Express $\mathbf{e}_1$ using the eigenvectors as a basis,

$$\mathbf{e}_1 = \alpha\hat{\mathbf{n}}_1 + \beta\hat{\mathbf{n}}_2 + \gamma\hat{\mathbf{n}}_3$$

Then

$$A'_{11} = \begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \alpha^2 \lambda_1 + \beta^2 \lambda_2 + \gamma^2 \lambda_3$$

Without loss of generality, let $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Then, with $\alpha^2 + \beta^2 + \gamma^2 = 1$, one has

$$\begin{aligned} \lambda_1 &= \lambda_1\left(\alpha^2 + \beta^2 + \gamma^2\right) \geq \alpha^2 \lambda_1 + \beta^2 \lambda_2 + \gamma^2 \lambda_3 = A'_{11} \\ \lambda_3 &= \lambda_3\left(\alpha^2 + \beta^2 + \gamma^2\right) \leq \alpha^2 \lambda_1 + \beta^2 \lambda_2 + \gamma^2 \lambda_3 = A'_{11} \end{aligned}$$

which proves that the eigenvalues include the largest and smallest possible diagonal element of **A**.

1.11.4 The Cayley-Hamilton Theorem

The **Cayley-Hamilton theorem** states that a tensor **A** (not necessarily symmetric) satisfies its own characteristic equation 1.11.5:

$$\mathbf{A}^3 - \mathrm{I}_{\mathbf{A}} \mathbf{A}^2 + \mathrm{II}_{\mathbf{A}} \mathbf{A} - \mathrm{III}_{\mathbf{A}} \mathbf{I} = \mathbf{0} \tag{1.11.15}$$

This can be proved as follows: one has $\mathbf{A}\hat{\mathbf{n}} = \lambda \hat{\mathbf{n}}$, where λ is an eigenvalue of **A** and $\hat{\mathbf{n}}$ is the corresponding eigenvector. A repeated application of **A** to this equation leads to $\mathbf{A}^n \hat{\mathbf{n}} = \lambda^n \hat{\mathbf{n}}$. Multiplying 1.11.5 by $\hat{\mathbf{n}}$ then leads to 1.11.15.

The third invariant in Eqn. 1.11.6 can now be written in terms of traces by a double contraction of the Cayley-Hamilton equation with **I**, and by using the definition of the trace, Eqn.1.10.6:

$$\begin{aligned} & \mathbf{A}^3 : \mathbf{I} - \mathrm{I}_{\mathbf{A}} \mathbf{A}^2 : \mathbf{I} + \mathrm{II}_{\mathbf{A}} \mathbf{A} : \mathbf{I} - \mathrm{III}_{\mathbf{A}} \mathbf{I} : \mathbf{I} = 0 \\ \rightarrow \quad & \mathrm{tr}\, \mathbf{A}^3 - \mathrm{I}_{\mathbf{A}} \mathrm{tr} \mathbf{A}^2 + \mathrm{II}_{\mathbf{A}} \mathrm{tr}\, \mathbf{A} - 3\mathrm{III}_{\mathbf{A}} = 0 \\ \rightarrow \quad & \mathrm{tr}\, \mathbf{A}^3 - \mathrm{tr}\, \mathbf{A} \mathrm{tr} \mathbf{A}^2 + \tfrac{1}{2}\left[(\mathrm{tr}\, \mathbf{A})^2 - \mathrm{tr}\, \mathbf{A}^2\right] \mathrm{tr}\, \mathbf{A} - 3\mathrm{III}_{\mathbf{A}} = 0 \\ \rightarrow \quad & \mathrm{III}_{\mathbf{A}} = \tfrac{1}{3}\left[\mathrm{tr}\, \mathbf{A}^3 - \tfrac{3}{2} \mathrm{tr}\, \mathbf{A}\, \mathrm{tr} \mathbf{A}^2 + \tfrac{1}{2}(\mathrm{tr}\, \mathbf{A})^3\right] \end{aligned} \tag{1.11.16}$$

The three invariants of a tensor can now be listed as

$$\boxed{\begin{aligned} \mathrm{I}_{\mathbf{A}} &= \mathrm{tr} \mathbf{A} \\ \mathrm{II}_{\mathbf{A}} &= \tfrac{1}{2}\left[(\mathrm{tr} \mathbf{A})^2 - \mathrm{tr}(\mathbf{A}^2)\right] \\ \mathrm{III}_{\mathbf{A}} &= \tfrac{1}{3}\left[\mathrm{tr}\, \mathbf{A}^3 - \tfrac{3}{2} \mathrm{tr}\, \mathbf{A}\, \mathrm{tr} \mathbf{A}^2 + \tfrac{1}{2}(\mathrm{tr}\, \mathbf{A})^3\right] \end{aligned}} \qquad \textbf{Invariants of a Tensor} \tag{1.11.17}$$

The Deviatoric Tensor

Denote the eigenvalues of the deviatoric tensor dev**A**, Eqn. 1.10.36, s_1, s_2, s_3 and the principal scalar invariants by J_1, J_2, J_3. The characteristic equation analogous to Eqn. 1.11.5 is then

$$s^3 - J_1 s^2 - J_2 s - J_3 = 0 \qquad (1.11.18)$$

and the deviatoric invariants are[3]

$$\begin{aligned} J_1 &= \mathrm{tr}(\mathrm{dev}\mathbf{A}) = s_1 + s_2 + s_3 \\ J_2 &= -\tfrac{1}{2}\left[(\mathrm{tr}(\mathrm{dev}\mathbf{A}))^2 - \mathrm{tr}\left((\mathrm{dev}\mathbf{A})^2\right)\right] = -(s_1 s_2 + s_2 s_3 + s_3 s_1) \\ J_3 &= \det(\mathrm{dev}\mathbf{A}) = s_1 s_2 s_3 \end{aligned} \qquad (1.11.19)$$

From Eqn. 1.10.37,

$$J_1 = 0 \qquad (1.11.20)$$

The second invariant can also be expressed in the useful forms {▲Problem 4}

$$J_2 = \tfrac{1}{2}\left(s_1^2 + s_2^2 + s_3^2\right), \qquad (1.11.21)$$

and, in terms of the eigenvalues of **A**, {▲Problem 5}

$$J_2 = \frac{1}{6}\left[(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2\right]. \qquad (1.11.22)$$

Further, the deviatoric invariants are related to the tensor invariants through {▲Problem 6}

$$J_2 = \tfrac{1}{3}\left(\mathrm{I}_{\mathbf{A}}^2 - 3\mathrm{II}_{\mathbf{A}}\right), \quad J_3 = \tfrac{1}{27}\left(2\mathrm{I}_{\mathbf{A}}^3 - 9\mathrm{I}_{\mathbf{A}}\mathrm{II}_{\mathbf{A}} + 27\mathrm{III}_{\mathbf{A}}\right) \qquad (1.11.23)$$

1.11.5 Coaxial Tensors

Two tensors are **coaxial** if they have the same eigenvectors. It can be shown that a necessary and sufficient condition that two tensors **A** and **B** be coaxial is that their simple contraction is commutative, $\mathbf{AB} = \mathbf{BA}$.

Since for a tensor **T**, $\mathbf{T}\mathbf{T}^{-1} = \mathbf{T}^{-1}\mathbf{T}$, a tensor and its inverse are coaxial and have the same eigenvectors.

---

[3] there is a convention (adhered to by most authors) to write the characteristic equation for a general tensor with a $+II_{\mathbf{A}}\lambda$ term and that for a deviatoric tensor with a $-J_2 s$ term (which ensures that $J_2 > 0$ - see 1.11.22 below) ; this means that the formulae for J_2 in Eqn. 1.11.19 are the negative of those for $II_{\mathbf{A}}$ in Eqn. 1.11.6

1.11.6 Fractional Powers of Tensors

Integer powers of tensors were defined in §1.9.2. Fractional powers of tensors can be defined provided the tensor is real, symmetric and positive definite (so that the eigenvalues are all positive).

Contracting both sides of $\mathbf{T}\hat{\mathbf{n}} = \lambda\hat{\mathbf{n}}$ with $\mathbf{T}$ repeatedly gives $\mathbf{T}^n\hat{\mathbf{n}} = \lambda^n\hat{\mathbf{n}}$. It follows that, if $\mathbf{T}$ has eigenvectors $\hat{\mathbf{n}}_i$ and corresponding eigenvalues λ_i, then $\mathbf{T}^n$ is coaxial, having the same eigenvectors, but corresponding eigenvalues λ_i^n. Because of this, fractional powers of tensors are defined as follows: $\mathbf{T}^m$, where *m* is any real number, is that tensor which has the same eigenvectors as $\mathbf{T}$ but which has corresponding eigenvalues λ_i^m. For example, the square root of the positive definite tensor $\mathbf{T} = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i$ is

$$\mathbf{T}^{1/2} = \sum_{i=1}^{3} \sqrt{\lambda_i}\, \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \qquad (1.11.24)$$

and the inverse is

$$\mathbf{T}^{-1} = \sum_{i=1}^{3} (1/\lambda_i)\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \qquad (1.11.25)$$

These new tensors are also positive definite.

1.11.7 Polar Decomposition of Tensors

Any (non-singular second-order) tensor $\mathbf{F}$ can be split up multiplicatively into an arbitrary proper orthogonal tensor $\mathbf{R}$ ($\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$, $\det \mathbf{R} = 1$) and a tensor $\mathbf{U}$ as follows:

$$\boxed{\mathbf{F} = \mathbf{R}\mathbf{U}} \qquad \textbf{Polar Decomposition} \qquad (1.11.26)$$

The consequence of this is that any transformation of a vector $\mathbf{a}$ according to $\mathbf{Fa}$ can be decomposed into two transformations, one involving a transformation $\mathbf{U}$, followed by a rotation $\mathbf{R}$.

The decomposition is not, in general, unique; one can often find more than one orthogonal tensor $\mathbf{R}$ which will satisfy the above relation. In practice, $\mathbf{R}$ is chosen such that $\mathbf{U}$ is symmetric. To this end, consider $\mathbf{F}^{\mathrm{T}}\mathbf{F}$. Since

$$\mathbf{v} \cdot \mathbf{F}^{\mathrm{T}}\mathbf{F}\mathbf{v} = \mathbf{F}\mathbf{v} \cdot \mathbf{F}\mathbf{v} = |\mathbf{F}\mathbf{v}|^2 > 0,$$

$\mathbf{F}^{\mathrm{T}}\mathbf{F}$ is positive definite. Further, $\mathbf{F}^{\mathrm{T}}\mathbf{F} \equiv F_{ji}F_{jk}$ is clearly symmetric, i.e. the same result is obtained upon an interchange of *i* and *k*. Thus the square-root of $\mathbf{F}^{\mathrm{T}}\mathbf{F}$ can be taken: let $\mathbf{U}$ in 1.11.26 be given by

$$\mathbf{U} = \left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right)^{1/2} \tag{1.11.27}$$

and **U** is also symmetric positive definite. Then, with 1.10.3e,

$$\begin{aligned}\mathbf{R}^{\mathrm{T}}\mathbf{R} &= \left(\mathbf{F}\mathbf{U}^{-1}\right)^{\mathrm{T}}\left(\mathbf{F}\mathbf{U}^{-1}\right)\\ &= \mathbf{U}^{-\mathrm{T}}\mathbf{F}^{\mathrm{T}}\mathbf{F}\mathbf{U}^{-1}\\ &= \mathbf{U}^{-\mathrm{T}}\mathbf{U}\mathbf{U}\mathbf{U}^{-1}\\ &= \mathbf{I}\end{aligned} \tag{1.11.28}$$

Thus if **U** is symmetric, **R** is orthogonal. Further, from (1.10.16a,b) and (1.100.18d), $\det\mathbf{U} = \det\mathbf{F}$ and $\det\mathbf{R} = \det\mathbf{F}/\det\mathbf{U} = 1$ so that **R** is proper orthogonal. It can also be proved that this decomposition is unique.

An alternative decomposition is given by

$$\mathbf{F} = \mathbf{V}\mathbf{R} \tag{1.11.29}$$

Again, this decomposition is unique and **R** is proper orthogonal, this time with

$$\mathbf{V} = \left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right)^{1/2} \tag{1.11.30}$$

1.11.8 Problems

1. Find the eigenvalues, (normalised) eigenvectors and principal invariants of
   $$\mathbf{T} = \mathbf{I} + \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1$$
2. Derive the spectral decomposition 1.11.11 by writing the identity tensor as $\mathbf{I} = \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i$, and writing $\mathbf{A} = \mathbf{A}\mathbf{I}$. [Hint: $\hat{\mathbf{n}}_i$ is an eigenvector.]
3. Derive the characteristic equation and Cayley-Hamilton equation for a 2-D space. Let **A** be a second order tensor with square root $\mathbf{S} = \sqrt{\mathbf{A}}$. By using the Cayley-Hamilton equation for **S**, and relating $\det\mathbf{S}$, $\mathrm{tr}\mathbf{S}$ to $\det\mathbf{A}$, $\mathrm{tr}\mathbf{A}$ through the corresponding eigenvalues, show that $\sqrt{\mathbf{A}} = \dfrac{\mathbf{A} + \sqrt{\det\mathbf{A}}\,\mathbf{I}}{\sqrt{\mathrm{tr}\mathbf{A} + 2\sqrt{\det\mathbf{A}}}}$.
4. The second invariant of a deviatoric tensor is given by Eqn. 1.11.19b,
   $$J_2 = -\left(s_1 s_2 + s_2 s_3 + s_3 s_1\right)$$
   By squaring the relation $J_1 = s_1 + s_2 + s_3 = 0$, derive Eqn. 1.11.21,
   $J_2 = \frac{1}{2}\left(s_1^2 + s_2^2 + s_3^2\right)$
5. Use Eqns. 1.11.21 (and your work from Problem 4) and the fact that $\lambda_1 - \lambda_2 = s_1 - s_2$, etc. to derive Eqn. 1.11.22.
6. Use the fact that $s_1 + s_2 + s_3 = 0$ to show that

$$\begin{aligned}
\mathrm{I}_{\mathbf{A}} &= 3\lambda_m \\
\mathrm{II}_{\mathbf{A}} &= (s_1 s_2 + s_2 s_3 + s_3 s_1) + 3\lambda_m^2 \\
\mathrm{III}_{\mathbf{A}} &= s_1 s_2 s_3 + \sigma_m (s_1 s_2 + s_2 s_3 + s_3 s_1) + \lambda_m^3
\end{aligned}$$

where $\lambda_m = \frac{1}{3} A_{ii}$. Hence derive Eqns. 1.11.23.

7. Consider the tensor

$$\mathbf{F} = \begin{bmatrix} 2 & -2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(a) Verify that the polar decomposition for **F** is $\mathbf{F} = \mathbf{RU}$ where

$$\mathbf{R} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 3/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 3/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(verify that **R** is proper orthogonal).

(b) Evaluate $\mathbf{Fa}, \mathbf{Fb}$, where $\mathbf{a} = [1,1,0]^{\mathrm{T}}$, $\mathbf{b} = [0,1,0]^{\mathrm{T}}$ by evaluating the individual transformations $\mathbf{Ua}, \mathbf{Ub}$ followed by $\mathbf{R}(\mathbf{Ua}), \mathbf{R}(\mathbf{Ub})$. Sketch the vectors and their images. Note how **R** rotates the vectors into their final positions. Why does **U** only stretch **a** but stretches *and* rotates **b**?

(c) Evaluate the eigenvalues λ_i and eigenvectors $\hat{\mathbf{n}}_i$ of the tensor $\mathbf{F}^{\mathrm{T}}\mathbf{F}$. Hence determine the spectral decomposition (diagonal matrix representation) of $\mathbf{F}^{\mathrm{T}}\mathbf{F}$. Hence evaluate $\mathbf{U} = \sqrt{\mathbf{F}^{\mathrm{T}}\mathbf{F}}$ with respect to the basis $\{\hat{\mathbf{n}}_i\}$ – again, this will be a diagonal matrix.

1.12 Higher Order Tensors

In this section are discussed some important higher (third and fourth) order tensors.

1.12.1 Fourth Order Tensors

After second-order tensors, the most commonly encountered tensors are the fourth order tensors $\mathbf{A}$, which have 81 components. Some properties and relations involving these tensors are listed here.

Transpose

The transpose of a fourth-order tensor $\mathbf{A}$, denoted by $\mathbf{A}^{\mathrm{T}}$, by analogy with the definition for the transpose of a second order tensor 1.10.4, is defined by

$$\mathbf{B} : \mathbf{A}^{\mathrm{T}} : \mathbf{C} = \mathbf{C} : \mathbf{A} : \mathbf{B} \tag{1.12.1}$$

for all second-order tensors **B** and **C**. It has the property $\left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathbf{A}$ and its components are $(\mathbf{A}^{\mathrm{T}})_{ijkl} = (\mathbf{A})_{klij}$. It also follows that

$$\left(\mathbf{A} \otimes \mathbf{B}\right)^{\mathrm{T}} = \mathbf{B} \otimes \mathbf{A} \tag{1.12.2}$$

Identity Tensors

There are two **fourth-order identity tensors**. They are defined as follows:

$$\begin{aligned} \mathbf{I} : \mathbf{A} &= \mathbf{A} \\ \bar{\mathbf{I}} : \mathbf{A} &= \mathbf{A}^{\mathrm{T}} \end{aligned} \tag{1.12.3}$$

And have components

$$\begin{aligned} \mathbf{I} &\equiv \delta_{ik}\delta_{jl}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l = \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_i \otimes \mathbf{e}_j \\ \bar{\mathbf{I}} &\equiv \delta_{il}\delta_{jk}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l = \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_j \otimes \mathbf{e}_i \end{aligned} \tag{1.12.4}$$

For a *symmetric* second order tensor **S**, $\bar{\mathbf{I}} : \mathbf{S} = \mathbf{I} : \mathbf{S} = \mathbf{S}$.

Another important fourth-order tensor is $\mathbf{I} \otimes \mathbf{I}$,

$$\mathbf{I} \otimes \mathbf{I} = \delta_{ij}\delta_{kl}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l = \mathbf{e}_i \otimes \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_j \tag{1.12.5}$$

Functions of the trace can be written in terms of these tensors {▲Problem 1}:

$$\begin{aligned}\mathbf{I}\otimes\mathbf{I}:\mathbf{A} &= (\mathrm{tr}\mathbf{A})\mathbf{I}\\ \mathbf{I}\otimes\mathbf{I}:\mathbf{A}:\mathbf{A} &= (\mathrm{tr}\mathbf{A})^2\\ \mathsf{I}:\mathbf{A}:\mathbf{A} &= \mathrm{tr}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)\\ \bar{\mathsf{I}}:\mathbf{A}:\mathbf{A} &= \mathrm{tr}\mathbf{A}^2\end{aligned} \qquad (1.12.6)$$

Projection Tensors

The symmetric and skew-symmetric parts of a second order tensor **A** can be written in terms of the identity tensors:

$$\begin{aligned}\mathrm{sym}\mathbf{A} &= \frac{1}{2}\left(\mathsf{I}+\bar{\mathsf{I}}\right):\mathbf{A}\\ \mathrm{skew}\mathbf{A} &= \frac{1}{2}\left(\mathsf{I}-\bar{\mathsf{I}}\right):\mathbf{A}\end{aligned} \qquad (1.12.7)$$

The deviator of **A**, 1.10.36, can be written as

$$\mathrm{dev}\mathbf{A} = \mathbf{A}-\frac{1}{3}(\mathrm{tr}\mathbf{A})\mathbf{I} = \mathbf{A}-\frac{1}{3}(\mathbf{I}:\mathbf{A})\mathbf{I} = \left(\mathsf{I}-\frac{1}{3}(\mathbf{I}\otimes\mathbf{I})\right):\mathbf{A}\equiv\hat{\mathsf{P}}:\mathbf{A} \qquad (1.12.8)$$

which defines $\hat{\mathsf{P}}$, the so-called **fourth-order projection tensor**. From Eqns. 1.10.6, 1.10.37a, it has the property that $\hat{\mathsf{P}}:\mathbf{A}:\mathbf{I}=0$. Note also that it has the property $\hat{\mathsf{P}}^n=\hat{\mathsf{P}}:\hat{\mathsf{P}}:\cdots:\hat{\mathsf{P}}=\hat{\mathsf{P}}$. For example,

$$\begin{aligned}\hat{\mathsf{P}}^2 = \hat{\mathsf{P}}:\hat{\mathsf{P}} &= \left(\mathsf{I}-\frac{1}{3}\mathbf{I}\otimes\mathbf{I}\right):\left(\mathsf{I}-\frac{1}{3}\mathbf{I}\otimes\mathbf{I}\right)\\ &= \mathsf{I}:\mathsf{I}-\frac{1}{3}\mathbf{I}\otimes\mathbf{I}-\frac{1}{3}\mathbf{I}\otimes\mathbf{I}+\frac{1}{9}(\mathbf{I}\otimes\mathbf{I}):(\mathbf{I}\otimes\mathbf{I})=\hat{\mathsf{P}}\end{aligned} \qquad (1.12.9)$$

The tensors $\left(\mathsf{I}+\bar{\mathsf{I}}\right)/2$, $\left(\mathsf{I}-\bar{\mathsf{I}}\right)/2$ in Eqn. 1.12.7 are also projection tensors, projecting the tensor **A** onto its symmetric and skew-symmetric parts.

1.12.2 Higher-Order Tensors and Symmetry

A higher order tensor possesses complete symmetry if the interchange of any indices is immaterial, for example if

$$\mathbf{A} = A_{ijk}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k) = A_{ikj}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k) = A_{jik}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k)=\cdots$$

It is symmetric in two of its indices if the interchange of these indices is immaterial. For example the above tensor $\mathbf{A}$ is symmetric in *j* and *k* if

$$\mathbf{A} = A_{ijk}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k) = A_{ikj}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k)$$

This applies also to antisymmetry. For example, the permutation tensor $\mathbf{E} = \varepsilon_{ijk}\left(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k\right)$ is completely antisymmetric, since $\varepsilon_{ijk} = -\varepsilon_{ikj} = \varepsilon_{kij} = \cdots$.

A fourth-order tensor $\mathbf{C}$ possesses the **minor symmetries** if

$$C_{ijkl} = C_{jikl}, \quad C_{ijkl} = C_{ijlk} \tag{1.12.10}$$

in which case it has only 36 independent components. The first equality here is for left minor symmetry, the second is for right minor symmetry.

It possesses the **major symmetries** if it *also* satisfies

$$C_{ijkl} = C_{klij} \tag{1.12.11}$$

in which case it has only 21 independent components. From 1.12.1, this can also be expressed as

$$\mathbf{A} : \mathbf{C} : \mathbf{B} = \mathbf{B} : \mathbf{C} : \mathbf{A} \tag{1.12.12}$$

for arbitrary second-order tensors **A**, **B**. Note that $\mathbf{I}, \bar{\mathbf{I}}, \mathbf{I} \otimes \mathbf{I}$ possess the major symmetries {▲Problem 2}.

1.12.3 Problems

1. Derive the relations 1.12.6.
2. Use 1.12.12 to show that $\mathbf{I}, \bar{\mathbf{I}}, \mathbf{I} \otimes \mathbf{I}$ possess the major symmetries.

1.13 Coordinate Transformation of Tensor Components

This section generalises the results of §1.5, which dealt with vector coordinate transformations. It has been seen in §1.5.2 that the transformation equations for the components of a vector are $u_i = Q_{ij} u'_j$, where $[\mathbf{Q}]$ is the transformation matrix. Note that these Q_{ij} 's are *not the components of a tensor* – these Q_{ij}'s are mapping the components of a vector onto the components of the *same vector* in a second coordinate system – a (second-order) tensor, in general, maps one vector onto a different vector. The equation $u_i = Q_{ij} u'_j$ is in matrix element form, and is not to be confused with the index notation for vectors and tensors.

1.13.1 Relationship between Base Vectors

Consider two coordinate systems with base vectors $\mathbf{e}_i$ and $\mathbf{e}'_i$. It has been seen in the context of vectors that, Eqn. 1.5.9,

$$\mathbf{e}_i \cdot \mathbf{e}'_j = Q_{ij} \equiv \cos(x_i, x'_j). \tag{1.13.1}$$

Recal that the *i*'s and *j*'s here are not referring to the three different components of a vector, but to *different* vectors (nine different vectors in all).

Note that the relationship 1.13.1 can also be derived as follows:

$$\begin{aligned} \mathbf{e}_i = \mathbf{I}\mathbf{e}_i &= (\mathbf{e}'_k \otimes \mathbf{e}'_k)\mathbf{e}_i \\ &= (\mathbf{e}'_k \cdot \mathbf{e}_i)\mathbf{e}'_k \\ &= Q_{ik}\mathbf{e}'_k \end{aligned} \tag{1.13.2}$$

Dotting each side here with $\mathbf{e}'_j$ then gives 1.13.1. Eqn. 1.13.2, together with the corresponding inverse relations, read

$$\mathbf{e}_i = Q_{ij}\mathbf{e}'_j, \qquad \mathbf{e}'_i = Q_{ji}\mathbf{e}_j \tag{1.13.3}$$

Note that the components of the transformation matrix $[\mathbf{Q}]$ are the same as the components of the change of basis tensor 1.10.24-25.

1.13.2 Tensor Transformation Rule

As with vectors, the components of a (second-order) tensor will change under a change of coordinate system. In this case, using 1.13.3,

$$\begin{aligned} T_{ij}\mathbf{e}_i \otimes \mathbf{e}_j &\equiv T'_{pq}\mathbf{e}'_p \otimes \mathbf{e}'_q \\ &= T'_{pq}Q_{mp}\mathbf{e}_m \otimes Q_{nq}\mathbf{e}_n \\ &= Q_{mp}Q_{nq}T'_{pq}\mathbf{e}_m \otimes \mathbf{e}_n \end{aligned} \tag{1.13.4}$$

so that (and the inverse relationship)

$$\boxed{T_{ij} = Q_{ip}Q_{jq}T'_{pq}, \quad T'_{ij} = Q_{pi}Q_{qj}T_{pq}} \quad \textbf{Tensor Transformation Formulae} \qquad (1.13.5)$$

or, in matrix form,

$$[\mathbf{T}] = [\mathbf{Q}][\mathbf{T}'][\mathbf{Q}^{\mathrm{T}}], \qquad [\mathbf{T}'] = [\mathbf{Q}^{\mathrm{T}}][\mathbf{T}][\mathbf{Q}] \qquad (1.13.6)$$

Note:
- as with vectors, second-order tensors are often *defined* as mathematical entities whose components transform according to the rule 1.13.5.
- the transformation rule for higher order tensors can be established in the same way, for example, $T'_{ijk} = Q_{pi}Q_{qj}Q_{rk}T_{pqr}$, and so on.

Example (Mohr Transformation)

Consider a two-dimensional space with base vectors $\mathbf{e}_1, \mathbf{e}_2$. The second order tensor **S** can be written in component form as

$$\mathbf{S} = S_{11}\mathbf{e}_1 \otimes \mathbf{e}_1 + S_{12}\mathbf{e}_1 \otimes \mathbf{e}_2 + S_{21}\mathbf{e}_2 \otimes \mathbf{e}_1 + S_{22}\mathbf{e}_2 \otimes \mathbf{e}_2$$

Consider now a second coordinate system, with base vectors $\mathbf{e}'_1, \mathbf{e}'_2$, obtained from the first by a rotation θ. The components of the transformation matrix are

$$Q_{ij} = \mathbf{e}_i \cdot \mathbf{e}'_j = \begin{bmatrix} \mathbf{e}_1 \cdot \mathbf{e}'_1 & \mathbf{e}_1 \cdot \mathbf{e}'_2 \\ \mathbf{e}_2 \cdot \mathbf{e}'_1 & \mathbf{e}_2 \cdot \mathbf{e}'_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \cos(90+\theta) \\ \cos(90-\theta) & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

and the components of **S** in the second coordinate system are $[\mathbf{S}'] = [\mathbf{Q}^{\mathrm{T}}][\mathbf{S}][\mathbf{Q}]$, so

$$\begin{bmatrix} S'_{11} & S'_{12} \\ S'_{21} & S'_{22} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

For **S** symmetric, $S_{12} = S_{21}$, and this simplifies to

$$\boxed{\begin{aligned} S'_{11} &= S_{11}\cos^2\theta + S_{22}\sin^2\theta + S_{12}\sin 2\theta \\ S'_{22} &= S_{11}\sin^2\theta + S_{22}\cos^2\theta - S_{12}\sin 2\theta \\ S'_{12} &= (S_{22} - S_{11})\sin\theta\cos\theta + S_{12}\cos 2\theta \end{aligned}} \quad \textbf{The Mohr Transformation} \qquad (1.13.7)$$

■

1.13.3 Isotropic Tensors

An **isotropic tensor** is one whose components are the same under arbitrary rotation of the basis vectors, i.e. in any coordinate system.

All scalars are isotropic.

There is no isotropic vector (first-order tensor), i.e. there is no vector **u** such that $u_i = Q_{ij} u_j$ for all orthogonal $[\mathbf{Q}]$ (except for the zero vector **o**). To see this, consider the particular orthogonal transformation matrix

$$[\mathbf{Q}] = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad (1.13.8)$$

which corresponds to a rotation of $\pi / 2$ about $\mathbf{e}_3$. This implies that

$$[u_1 \quad u_2 \quad u_3]^{\mathrm{T}} = [u_2 \quad -u_1 \quad u_3]^{\mathrm{T}}$$

or $u_1 = u_2 = 0$. The matrix corresponding to a rotation of $\pi / 2$ about $\mathbf{e}_1$ is

$$[\mathbf{Q}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad (1.13.9)$$

which implies that $u_3 = 0$.

The only isotropic second-order tensor is $\alpha \mathbf{I} \equiv \alpha \delta_{ij}$, where α is a constant, that is, the spherical tensor, §1.10.12. To see this, first note that, by substituting $\alpha \mathbf{I}$ into 1.13.6, it can be seen that it is indeed isotropic. To see that it is the only isotropic second order tensor, first use 1.13.8 in 1.13.6 to get

$$[\mathbf{T}'] = \begin{bmatrix} T_{22} & -T_{21} & -T_{23} \\ -T_{12} & T_{11} & T_{13} \\ -T_{32} & T_{31} & T_{33} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \qquad (1.13.10)$$

which implies that $T_{11} = T_{22}, T_{12} = -T_{21}, T_{13} = T_{23} = T_{31} = T_{32} = 0$. Repeating this for 1.13.9 implies that $T_{11} = T_{33}, T_{12} = 0$, so

$$[\mathbf{T}] = \begin{bmatrix} T_{11} & 0 & 0 \\ 0 & T_{11} & 0 \\ 0 & 0 & T_{11} \end{bmatrix}$$

or $\mathbf{T} = T_{11} \mathbf{I}$. Multiplying by a scalar does not affect 1.13.6, so one has $\alpha \mathbf{I}$.

The only third-order isotropic tensors are scalar multiples of the permutation tensor, $\mathbf{E} = \varepsilon_{ijk} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)$. Using the third order transformation rule, $T'_{ijk} = Q_{pi} Q_{qj} Q_{rk} T_{pqr}$, one has $\varepsilon'_{ijk} = Q_{pi} Q_{qj} Q_{rk} \varepsilon_{pqr}$. From 1.10.16e this reads $\varepsilon'_{ijk} = (\det \mathbf{Q}) \varepsilon_{ijk}$, where **Q** is the change of basis tensor, with components Q_{ij}. When **Q** is proper orthogonal, i.e. a rotation tensor, one has indeed, $\varepsilon'_{ijk} = \varepsilon_{ijk}$. That it is the only isotropic tensor can be established by carrying out a few specific rotations as done above for the first and second order tensors.

Note that orthogonal tensors in general, i.e. having the possibility of being reflection tensors, with $\det \mathbf{Q} = -1$, are not used in the definition of isotropy, otherwise one would have the less desirable $\varepsilon'_{ijk} = -\varepsilon_{ijk}$. Note also that this issue does not arise with the second order tensor (or the fourth order tensor –see below), since the above result, that $\alpha \mathbf{I}$ is the only isotropic second order tensor, holds regardless of whether **Q** is proper orthogonal or not.

There are three independent fourth-order isotropic tensors – these are the tensors encountered in §1.12.1, Eqns. 1.12.4-5,

$$\boldsymbol{\mathsf{I}}, \quad \bar{\boldsymbol{\mathsf{I}}}, \quad \mathbf{I} \otimes \mathbf{I}$$

For example,

$$Q_{ip} Q_{jq} Q_{kr} Q_{ls} (\mathbf{I} \otimes \mathbf{I})_{pqrs} = Q_{ip} Q_{jq} Q_{kr} Q_{ls} \delta_{pq} \delta_{rs} = (Q_{ip} Q_{jp})(Q_{kr} Q_{lr}) = \delta_{ij} \delta_{kl} = (\mathbf{I} \otimes \mathbf{I})_{ijkl}$$

The most general isotropic fourth order tensor is then a linear combination of these tensors:

$$\boxed{\boldsymbol{\mathsf{C}} = \lambda \mathbf{I} \otimes \mathbf{I} + \mu \boldsymbol{\mathsf{I}} + \gamma \bar{\boldsymbol{\mathsf{I}}}} \quad \textbf{Most General Isotropic Fourth-Order Tensor} \qquad (1.13.11)$$

1.13.4 Invariance of Tensor Components

The components of (non-isotropic) tensors will change upon a rotation of base vectors. However, certain combinations of these components are the same in *every* coordinate system. Such quantities are called **invariants**. For example, the following are examples of **scalar invariants** {▲Problem 2}

$$\begin{aligned} \mathbf{a} \cdot \mathbf{a} &= a_i a_i \\ \mathbf{a} \cdot \mathbf{T} \mathbf{a} &= T_{ij} a_i a_j \\ \operatorname{tr} \mathbf{A} &= A_{ii} \end{aligned} \qquad (1.13.12)$$

The first of these is the only independent scalar invariant of a vector. A second-order tensor has three independent scalar invariants, the first, second and third principal scalar invariants, defined by Eqn. 1.11.17 (or linear combinations of these).

1.13.5 Problems

1. Consider a coordinate system $ox_1x_2x_3$ with base vectors $\mathbf{e}_i$. Let a second coordinate system be represented by the set $\{\mathbf{e}'_i\}$ with the transformation law

$$\mathbf{e}'_2 = -\sin\theta\mathbf{e}_1 + \cos\theta\mathbf{e}_2$$
$$\mathbf{e}'_3 = \mathbf{e}_3$$

   (a) find $\mathbf{e}'_1$ in terms of the old set $\{\mathbf{e}_i\}$ of basis vectors
   (b) find the orthogonal matrix $[\mathbf{Q}]$ and express the old coordinates in terms of the new ones
   (c) express the vector $\mathbf{u} = -6\mathbf{e}_1 - 3\mathbf{e}_2 + \mathbf{e}_3$ in terms of the new set $\{\mathbf{e}'_i\}$ of basis vectors.
2. Show that
   (a) the trace of a tensor **A**, $\mathrm{tr}\mathbf{A} = A_{ii}$, is an invariant.
   (b) $\mathbf{a} \cdot \mathbf{T}\mathbf{a} = T_{ij}a_i a_j$ is an invariant.
3. Consider Problem 7 in §1.11. Take the tensor $\mathbf{U} = \sqrt{\mathbf{F}^{\mathrm{T}}\mathbf{F}}$ with respect to the basis $\{\hat{\mathbf{n}}_i\}$ and carry out a coordinate transformation of its tensor components so that it is given with respect to the original $\{\mathbf{e}_i\}$ basis – in which case the matrix representation for **U** given in Problem 7, §1.11, should be obtained.

1.14 Tensor Calculus I: Tensor Fields

In this section, the concepts from the calculus of vectors are generalised to the calculus of higher-order tensors.

1.14.1 Tensor-valued Functions

Tensor-valued functions of a scalar

The most basic type of calculus is that of tensor-valued functions of a scalar, for example the time-dependent stress at a point, $\mathbf{S} = \mathbf{S}(t)$. If a tensor $\mathbf{T}$ depends on a scalar t, then the derivative is defined in the usual way,

$$\frac{d\mathbf{T}}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{T}(t + \Delta t) - \mathbf{T}(t)}{\Delta t},$$

which turns out to be

$$\frac{d\mathbf{T}}{dt} = \frac{dT_{ij}}{dt} \mathbf{e}_i \otimes \mathbf{e}_j \tag{1.14.1}$$

The derivative is also a tensor and the usual rules of differentiation apply,

$$\begin{aligned}
\frac{d}{dt}(\mathbf{T} + \mathbf{B}) &= \frac{d\mathbf{T}}{dt} + \frac{d\mathbf{B}}{dt} \\
\frac{d}{dt}(\alpha(t)\mathbf{T}) &= \alpha \frac{d\mathbf{T}}{dt} + \frac{d\alpha}{dt}\mathbf{T} \\
\frac{d}{dt}(\mathbf{T}\mathbf{a}) &= \mathbf{T}\frac{d\mathbf{a}}{dt} + \frac{d\mathbf{T}}{dt}\mathbf{a} \\
\frac{d}{dt}(\mathbf{T}\mathbf{B}) &= \mathbf{T}\frac{d\mathbf{B}}{dt} + \frac{d\mathbf{T}}{dt}\mathbf{B} \\
\frac{d}{dt}\left(\mathbf{T}^{\mathrm{T}}\right) &= \left(\frac{d\mathbf{T}}{dt}\right)^{\mathrm{T}}
\end{aligned}$$

For example, consider the time derivative of $\mathbf{Q}\mathbf{Q}^{\mathrm{T}}$, where $\mathbf{Q}$ is orthogonal. By the product rule, using $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$,

$$\frac{d}{dt}\left(\mathbf{Q}\mathbf{Q}^{\mathrm{T}}\right) = \frac{d\mathbf{Q}}{dt}\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}\frac{d\mathbf{Q}^{\mathrm{T}}}{dt} = \frac{d\mathbf{Q}}{dt}\mathbf{Q}^{\mathrm{T}} + \mathbf{Q}\left(\frac{d\mathbf{Q}}{dt}\right)^{\mathrm{T}} = \mathbf{0}$$

Thus, using Eqn. 1.10.3e

$$\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\dot{\mathbf{Q}}^{\mathrm{T}} = -\left(\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}\right)^{\mathrm{T}} \tag{1.14.2}$$

which shows that $\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}$ is a skew-symmetric tensor.

1.14.2 Vector Fields

The gradient of a scalar field and the divergence and curl of vector fields have been seen in §1.6. Other important quantities are the gradient of vectors and higher order tensors and the divergence of higher order tensors. First, the gradient of a vector field is introduced.

The Gradient of a Vector Field

The gradient of a vector field is defined to be the second-order tensor

$$\boxed{\mathrm{grad}\mathbf{a} \equiv \frac{\partial \mathbf{a}}{\partial x_j} \otimes \mathbf{e}_j = \frac{\partial a_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j} \quad \textbf{Gradient of a Vector Field} \qquad (1.14.3)$$

In matrix notation,

$$\mathrm{grad}\mathbf{a} = \begin{bmatrix} \frac{\partial a_1}{\partial x_1} & \frac{\partial a_1}{\partial x_2} & \frac{\partial a_1}{\partial x_3} \\ \frac{\partial a_2}{\partial x_1} & \frac{\partial a_2}{\partial x_2} & \frac{\partial a_2}{\partial x_3} \\ \frac{\partial a_3}{\partial x_1} & \frac{\partial a_3}{\partial x_2} & \frac{\partial a_3}{\partial x_3} \end{bmatrix} \qquad (1.14.4)$$

One then has

$$\begin{aligned} \mathrm{grad}\mathbf{a}\, d\mathbf{x} &= \frac{\partial a_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j \left(dx_k \mathbf{e}_k \right) \\ &= \frac{\partial a_i}{\partial x_j} dx_j \mathbf{e}_i \\ &= d\mathbf{a} \\ &= \mathbf{a}(\mathbf{x} + d\mathbf{x}) - \mathbf{a}(d\mathbf{x}) \end{aligned} \qquad (1.14.5)$$

which is analogous to Eqn 1.6.10 for the gradient of a scalar field. As with the gradient of a scalar field, if one writes $d\mathbf{x}$ as $|d\mathbf{x}|\mathbf{e}$, where $\mathbf{e}$ is a unit vector, then

$$\mathrm{grad}\mathbf{a}\, \mathbf{e} = \left(\frac{d\mathbf{a}}{dx} \right)_{\text{in } \mathbf{e} \text{ direction}} \qquad (1.14.6)$$

Thus the gradient of a vector field $\mathbf{a}$ is a second-order tensor which transforms a unit vector into a vector describing the gradient of $\mathbf{a}$ in that direction.

As an example, consider a space curve parameterised by s, with unit tangent vector $\boldsymbol{\tau} = d\mathbf{x} / ds$ (see §1.6.2); one has

$$\frac{d\mathbf{a}}{ds} = \frac{\partial \mathbf{a}}{\partial x_j}\frac{dx_j}{ds} = \frac{\partial \mathbf{a}}{\partial x_j}\left(\boldsymbol{\tau} \cdot \mathbf{e}_j\right) = \left(\frac{\partial \mathbf{a}}{\partial x_j} \otimes \mathbf{e}_j\right)\boldsymbol{\tau} = \operatorname{grad}\mathbf{a}\,\boldsymbol{\tau}\,.$$

Although for a scalar field $\operatorname{grad}\phi$ is equivalent to $\nabla\phi$, note that the gradient defined in 1.14.3 is *not* the same as $\nabla \otimes \mathbf{a}$. In fact,

$$\left(\nabla \otimes \mathbf{a}\right)^{\mathrm{T}} = \operatorname{grad}\mathbf{a} \tag{1.14.7}$$

since

$$\nabla \otimes \mathbf{a} = \mathbf{e}_i \frac{\partial}{\partial x_i} \otimes a_j \mathbf{e}_j = \frac{\partial a_j}{\partial x_i}\mathbf{e}_i \otimes \mathbf{e}_j \tag{1.14.8}$$

These two different definitions of the gradient of a vector, $\partial a_i / \partial x_j \mathbf{e}_i \otimes \mathbf{e}_j$ and $\partial a_j / \partial x_i \mathbf{e}_i \otimes \mathbf{e}_j$, are both commonly used. In what follows, they will be distinguished by labeling the former as $\operatorname{grad}\mathbf{a}$ (which will be called the gradient of **a**) and the latter as $\nabla \otimes \mathbf{a}$.

Note the following:
- in much of the literature, $\nabla \otimes \mathbf{a}$ is written in the contracted form $\nabla\mathbf{a}$, but the more explicit version is used here.
- some authors define the operation of $\nabla \otimes$ on a vector or tensor $(\bullet)$ not as in 1.14.8, but through $\nabla \otimes (\bullet) \equiv \left(\partial(\bullet) / \partial x_i\right) \otimes \mathbf{e}_i$ so that $\nabla \otimes \mathbf{a} = \operatorname{grad}\mathbf{a} = \left(\partial a_i / \partial x_j\right)\mathbf{e}_i \otimes \mathbf{e}_j$.

Example (The Displacement Gradient)

Consider a particle p_0 of a deforming body at position $\mathbf{X}$ (a vector) and a neighbouring point q_0 at position $d\mathbf{X}$ relative to p_0, Fig. 1.14.1. As the material deforms, these two particles undergo displacements of, respectively, $\mathbf{u}(\mathbf{X})$ and $\mathbf{u}(\mathbf{X} + d\mathbf{X})$. The final positions of the particles are p_f and q_f. Then

$$\begin{aligned} d\mathbf{x} &= d\mathbf{X} + \mathbf{u}(\mathbf{X} + d\mathbf{X}) - \mathbf{u}(\mathbf{X}) \\ &= d\mathbf{X} + d\mathbf{u}(\mathbf{X}) \\ &= d\mathbf{X} + \operatorname{grad}\mathbf{u}\, d\mathbf{X} \end{aligned}$$

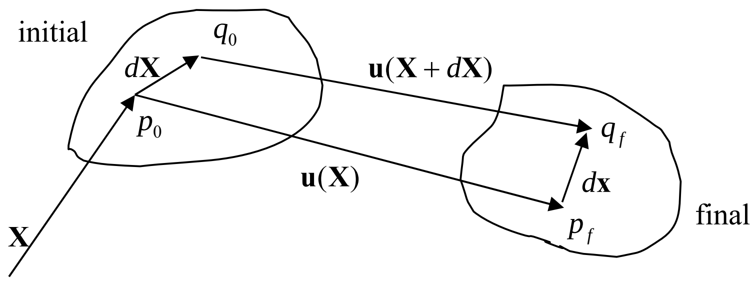


Figure 1.14.1: displacement of material particles

Thus the gradient of the displacement field **u** encompasses the mapping of (infinitesimal) line elements in the undeformed body into line elements in the deformed body. For example, suppose that $u_1 = kX_2^2, u_2 = u_3 = 0$. Then

$$\mathrm{grad}\mathbf{u} = \frac{\partial u_i}{\partial X_j} = \begin{bmatrix} 0 & 2kX_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 2kX_2 \mathbf{e}_1 \otimes \mathbf{e}_2$$

A line element $d\mathbf{X} = dX_i \mathbf{e}_i$ at $\mathbf{X} = X_i \mathbf{e}_i$ maps onto

$$\begin{aligned} d\mathbf{x} &= d\mathbf{X} + (2kX_2 \mathbf{e}_1 \otimes \mathbf{e}_2)(dX_1 \mathbf{e}_1 + dX_2 \mathbf{e}_2 + dX_3 \mathbf{e}_3) \\ &= d\mathbf{X} + 2kX_2 dX_2 \mathbf{e}_1 \end{aligned}$$

The deformation of a box is as shown in Fig. 1.14.2. For example, the vector $d\mathbf{X} = d\alpha \mathbf{e}_2$ (defining the left-hand side of the box) maps onto $d\mathbf{x} = 2k\alpha\, d\alpha \mathbf{e}_1 + \alpha \mathbf{e}_2$.

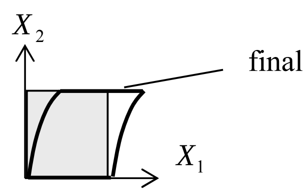


Figure 1.14.2: deformation of a box

Note that the map $d\mathbf{X} \to d\mathbf{x}$ does not specify where in space the line element moves to. It translates too according to $\mathbf{x} = \mathbf{X} + \mathbf{u}$.

■

The Divergence and Curl of a Vector Field

The divergence and curl of vectors have been defined in §1.6.6, §1.6.8. Now that the gradient of a vector has been introduced, one can re-define the divergence of a vector independent of any coordinate system: it is the scalar field given by the trace of the gradient {▲Problem 4},

$$\boxed{\operatorname{div}\mathbf{a} = \operatorname{tr}(\operatorname{grad}\mathbf{a}) = \operatorname{grad}\mathbf{a} : \mathbf{I} = \nabla \cdot \mathbf{a}} \quad \textbf{Divergence of a Vector Field} \qquad (1.14.9)$$

Similarly, the curl of **a** can be defined to be the vector field given by twice the axial vector of the antisymmetric part of $\operatorname{grad}\mathbf{a}$.

1.14.3 Tensor Fields

A tensor-valued function of the position vector is called a tensor field, $T_{ij\cdots k}(\mathbf{x})$.

The Gradient of a Tensor Field

The gradient of a second order tensor field **T** is defined in a manner analogous to that of the gradient of a vector, Eqn. 1.14.2. It is the third-order tensor

$$\boxed{\operatorname{grad}\mathbf{T} = \frac{\partial \mathbf{T}}{\partial x_k} \otimes \mathbf{e}_k = \frac{\partial T_{ij}}{\partial x_k} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k} \quad \textbf{Gradient of a Tensor Field} \qquad (1.14.10)$$

This differs from the quantity

$$\nabla \otimes \mathbf{T} = \mathbf{e}_i \frac{\partial}{\partial x_i} \otimes \left(T_{jk} \mathbf{e}_j \otimes \mathbf{e}_k \right) = \frac{\partial T_{jk}}{\partial x_i} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \qquad (1.14.11)$$

The Divergence of a Tensor Field

Analogous to the definition 1.14.9, the divergence of a second order tensor **T** is defined to be the vector

$$\boxed{\begin{aligned} \operatorname{div}\mathbf{T} &= \operatorname{grad}\mathbf{T} : \mathbf{I} = \frac{\partial \mathbf{T}}{\partial x_i} \mathbf{e}_i = \frac{\partial (T_{jk} \mathbf{e}_j \otimes \mathbf{e}_k)}{\partial x_i} \mathbf{e}_i \\ &= \frac{\partial T_{ij}}{\partial x_j} \mathbf{e}_i \end{aligned}} \quad \textbf{Divergence of a Tensor} \qquad (1.14.12)$$

One also has

$$\nabla \cdot \mathbf{T} = \mathbf{e}_i \frac{\partial}{\partial x_i} \cdot (T_{jk} \mathbf{e}_j \otimes \mathbf{e}_k) = \frac{\partial T_{ji}}{\partial x_j} \mathbf{e}_i \qquad (1.14.13)$$

so that

$$\operatorname{div}\mathbf{T} = \nabla \cdot \mathbf{T}^{\mathrm{T}} \qquad (1.14.14)$$

As with the gradient of a vector, both $\left(\partial T_{ij} / \partial x_j\right)\mathbf{e}_i$ and $\left(\partial T_{ji} / \partial x_j\right)\mathbf{e}_i$ are commonly used as definitions of the divergence of a tensor. They are distinguished here by labelling the former as $\text{div}\mathbf{T}$ (called here the divergence of **T**) and the latter as $\nabla \cdot \mathbf{T}$. Note that the operations $\text{div}\mathbf{T}$ and $\nabla \cdot \mathbf{T}$ are equivalent for the case of **T** symmetric.

The Laplacian of a scalar θ is the scalar $\nabla^2\theta \equiv \nabla \cdot \nabla\theta$, in component form $\partial^2\theta / \partial x_i^2$ (see section 1.6.7). Similarly, the Laplacian of a vector **v** is the vector $\nabla^2\mathbf{v} \equiv \nabla \cdot \nabla\mathbf{v}$, in component form $\partial^2 v_i / \partial x_j^2$. The Laplacian of a tensor **T** in component form is similarly $\partial^2 T_{ij} / \partial x_k^2$, which can be defined as that tensor field which satisfies the relation

$$\left(\nabla^2\mathbf{T}\right) \cdot \mathbf{v} = \nabla^2\left(\mathbf{T}\mathbf{v}\right)$$

for all constant vectors **v**.

Note the following

- some authors define the operation of $\nabla \cdot$ on a vector or tensor $(\bullet)$ not as in (1.14.13), but through $\nabla \cdot (\bullet) \equiv \left(\partial(\bullet) / \partial x_i\right) \cdot \mathbf{e}_i$ so that $\nabla \cdot \mathbf{T} = \text{div}\mathbf{T} = \left(\partial T_{ij} / \partial x_j\right)\mathbf{e}_i$.
- using the convention that the "dot" is omitted in the contraction of tensors, one should write $\nabla\mathbf{T}$ for $\nabla \cdot \mathbf{T}$, but the "dot" is retained here because of the familiarity of this latter notation from vector calculus.
- another operator is the **Hessian**, $\nabla \otimes \nabla = \left(\partial^2 / \partial x_i \partial x_j\right)\mathbf{e}_i \otimes \mathbf{e}_j$.

Identities

Here are some important identities involving the gradient, divergence and curl {▲Problem 5}:

$$\begin{aligned} \text{grad}(\phi\mathbf{v}) &= \phi\,\text{grad}\mathbf{v} + \mathbf{v} \otimes \text{grad}\phi \\ \text{grad}(\mathbf{u} \cdot \mathbf{v}) &= (\text{grad}\mathbf{u})^{\mathrm{T}}\,\mathbf{v} + (\text{grad}\mathbf{v})^{\mathrm{T}}\,\mathbf{u} \\ \text{div}(\mathbf{u} \otimes \mathbf{v}) &= (\text{grad}\,\mathbf{u})\mathbf{v} + (\text{div}\,\mathbf{v})\mathbf{u} \\ \text{curl}(\mathbf{u} \times \mathbf{v}) &= \mathbf{u}\,\text{div}\mathbf{v} - \mathbf{v}\,\text{div}\mathbf{u} + (\text{grad}\mathbf{u})\mathbf{v} - (\text{grad}\mathbf{v})\mathbf{u} \end{aligned} \qquad (1.14.15)$$

$$\begin{aligned} \text{div}(\phi\mathbf{A}) &= \mathbf{A}\,\text{grad}\phi + \phi\,\text{div}\mathbf{A} \\ \text{div}(\mathbf{A}\mathbf{v}) &= \mathbf{v} \cdot \text{div}\mathbf{A}^{\mathrm{T}} + \text{tr}\left(\mathbf{A}\,\text{grad}\mathbf{v}\right) \\ \text{div}(\mathbf{A}\mathbf{B}) &= \mathbf{A}\,\text{div}\mathbf{B} + \text{grad}\mathbf{A} : \mathbf{B} \\ \text{div}(\mathbf{A}(\phi\mathbf{B})) &= \phi\,\text{div}(\mathbf{A}\mathbf{B}) + \mathbf{A}(\mathbf{B}\,\text{grad}\phi) \\ \text{grad}(\phi\mathbf{A}) &= \phi\,\text{grad}\mathbf{A} + \mathbf{A} \otimes \text{grad}\phi \end{aligned} \qquad (1.11.16)$$

Note also the following identities, which involve the Laplacian of both vectors and scalars:

$$\begin{aligned} \nabla^2(\mathbf{u} \cdot \mathbf{v}) &= \nabla^2\mathbf{u} \cdot \mathbf{v} + 2\,\text{grad}\mathbf{u} : \text{grad}\mathbf{v} + \mathbf{u} \cdot \nabla^2\mathbf{v} \\ \text{curl}\,\text{curl}\,\mathbf{u} &= \text{grad}(\text{div}\,\mathbf{u}) - \nabla^2\mathbf{u} \end{aligned} \qquad (1.14.17)$$

1.14.4 Cylindrical and Spherical Coordinates

Cylindrical and spherical coordinates were introduced in §1.6.10 and the gradient and Laplacian of a scalar field and the divergence and curl of vector fields were derived in terms of these coordinates. The calculus of higher order tensors can also be cast in terms of these coordinates.

For example, from 1.6.30, the gradient of a vector in cylindrical coordinates is $\mathrm{grad}\mathbf{u} = (\nabla \otimes \mathbf{u})^{\mathrm{T}}$ with

$$\begin{aligned}
\mathrm{grad}\mathbf{u} &= \left[\left(\mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \mathbf{e}_z \frac{\partial}{\partial z}\right) \otimes \left(u_r \mathbf{e}_r + u_\theta \mathbf{e}_\theta + u_z \mathbf{e}_z\right)\right]^{\mathrm{T}} \\
&= \frac{\partial u_r}{\partial r}\mathbf{e}_r \otimes \mathbf{e}_r + \left(\frac{1}{r}\frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r}\right)\mathbf{e}_r \otimes \mathbf{e}_\theta + \frac{\partial u_r}{\partial z}\mathbf{e}_r \otimes \mathbf{e}_z \\
&\quad + \frac{\partial u_\theta}{\partial r}\mathbf{e}_\theta \otimes \mathbf{e}_r + \left(\frac{1}{r}\frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r}\right)\mathbf{e}_\theta \otimes \mathbf{e}_\theta + \frac{\partial u_\theta}{\partial z}\mathbf{e}_\theta \otimes \mathbf{e}_z \\
&\quad + \frac{\partial u_z}{\partial r}\mathbf{e}_z \otimes \mathbf{e}_r + \frac{1}{r}\frac{\partial u_z}{\partial \theta}\mathbf{e}_z \otimes \mathbf{e}_\theta + \frac{\partial u_z}{\partial z}\mathbf{e}_z \otimes \mathbf{e}_z
\end{aligned} \tag{1.14.18}$$

and from 1.6.30, 1.14.12, the divergence of a tensor in cylindrical coordinates is {▲Problem 6}

$$\begin{aligned}
\mathrm{div}\mathbf{A} = \nabla \cdot \mathbf{A}^{\mathrm{T}} &= \left(\frac{\partial A_{rr}}{\partial r} + \frac{1}{r}\frac{\partial A_{r\theta}}{\partial \theta} + \frac{\partial A_{rz}}{\partial z} + \frac{A_{rr} - A_{\theta\theta}}{r}\right)\mathbf{e}_r \\
&\quad + \left(\frac{\partial A_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial A_{\theta\theta}}{\partial \theta} + \frac{\partial A_{\theta z}}{\partial z} + \frac{A_{\theta r} + A_{r\theta}}{r}\right)\mathbf{e}_\theta \\
&\quad + \left(\frac{\partial A_{zr}}{\partial r} + \frac{A_{zr}}{r} + \frac{1}{r}\frac{\partial A_{z\theta}}{\partial \theta} + \frac{\partial A_{zz}}{\partial z}\right)\mathbf{e}_z
\end{aligned} \tag{1.14.19}$$

1.14.5 The Divergence Theorem

The divergence theorem 1.7.12 can be extended to the case of higher-order tensors. Consider an arbitrary differentiable tensor field $T_{ij\cdots k}(\mathbf{x},t)$ defined in some finite region of physical space. Let S be a closed surface bounding a volume V in this space, and let the outward normal to S be $\mathbf{n}$. The divergence theorem of Gauss then states that

$$\int_S T_{ij\cdots k} n_k \, dS = \int_V \frac{\partial T_{ij\cdots k}}{\partial x_k} dV \tag{1.14.20}$$

For a second order tensor,

$$\int_S \mathbf{T}\mathbf{n}\, dS = \int_V \operatorname{div} \mathbf{T}\, dV, \qquad \int_S T_{ij} n_j\, dS = \int_V \frac{\partial T_{ij}}{\partial x_j} dV \tag{1.14.21}$$

One then has the important identities {▲Problem 7}

$$\begin{aligned} \int_S (\phi \mathbf{T})\mathbf{n}\, dS &= \int_V \operatorname{div}(\phi \mathbf{T})\, dV \\ \int_S \mathbf{u} \otimes \mathbf{n}\, dS &= \int_V \operatorname{grad} \mathbf{u}\, dV \\ \int_S \mathbf{u} \cdot \mathbf{T}\mathbf{n}\, dS &= \int_V \operatorname{div}(\mathbf{T}^{\mathrm{T}}\mathbf{u})\, dV \end{aligned} \tag{1.14.22}$$

1.14.6 Formal Treatment of Tensor Calculus

Following on from §1.6.12, here a more formal treatment of the tensor calculus of fields is briefly presented.

Vector Gradient

What follows is completely analogous to Eqns. 1.6.46-49.

A **vector field** $\mathbf{v} : E^3 \to V$ is **differentiable** at a point $\mathbf{x} \in E^3$ if there exists a second order tensor $D\mathbf{v}(\mathbf{x}) \in E$ such that

$$\mathbf{v}(\mathbf{x} + \mathbf{h}) = \mathbf{v}(\mathbf{x}) + D\mathbf{v}(\mathbf{x})\mathbf{h} + o(\|\mathbf{h}\|) \quad \text{for all} \quad \mathbf{h} \in E \tag{1.14.23}$$

In that case, the tensor $D\mathbf{v}(\mathbf{x})$ is called the **derivative** (or **gradient**) of **v** at **x** (and is given the symbol $\nabla\mathbf{v}(\mathbf{x})$).

Setting $\mathbf{h} = \varepsilon\mathbf{w}$ in 1.14.23, where $\mathbf{w} \in E$ is a unit vector, dividing through by ε and taking the limit as $\varepsilon \to 0$, one has the equivalent statement

$$\nabla\mathbf{v}(\mathbf{x})\mathbf{w} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{v}(\mathbf{x} + \varepsilon\mathbf{w}) \quad \text{for all} \quad \mathbf{w} \in E \tag{1.14.24}$$

Using the chain rule as in §1.6.11, Eqn. 1.14.24 can be expressed in terms of the Cartesian basis $\{\mathbf{e}_i\}$,

$$\nabla\mathbf{v}(\mathbf{x})\, \mathbf{w} = \frac{\partial v_i}{\partial x_k} w_k \mathbf{e}_i = \frac{\partial v_i}{\partial x_j}(\mathbf{e}_i \otimes \mathbf{e}_j) w_k \mathbf{e}_k \tag{1.14.25}$$

This must be true for all **w** and so, in a Cartesian basis,

$$\nabla \mathbf{v}(\mathbf{x}) = \frac{\partial v_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j \qquad (1.14.26)$$

which is Eqn. 1.14.3.

1.14.7 Problems

1. Consider the vector field $\mathbf{v} = x_1^2 \mathbf{e}_1 + x_3^2 \mathbf{e}_2 + x_2^2 \mathbf{e}_3$. (a) find the matrix representation of the gradient of **v**, (b) find the vector $(\mathrm{grad}\mathbf{v})\mathbf{v}$.
2. If $\mathbf{u} = x_1 x_2 x_3 \mathbf{e}_1 + x_1 x_2 \mathbf{e}_2 + x_1 \mathbf{e}_3$, determine $\nabla^2 \mathbf{u}$.
3. Suppose that the displacement field is given by $u_1 = 0, u_2 = 1, u_3 = X_1$. By using $\mathrm{grad}\mathbf{u}$, sketch a few (undeformed) line elements of material and their positions in the deformed configuration.
4. Use the matrix form of $\mathrm{grad}\mathbf{u}$ and $\nabla \otimes \mathbf{u}$ to show that the definitions
   (i) $\mathrm{div}\mathbf{a} = \mathrm{tr}(\mathrm{grad}\mathbf{a})$
   (ii) $\mathrm{curl}\mathbf{a} = 2\boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is the axial vector of the skew part of $\mathrm{grad}\mathbf{a}$
   agree with the definitions 1.6.17, 1.6.21 given for Cartesian coordinates.
5. Prove the following:
   (i) $\mathrm{grad}(\phi\mathbf{v}) = \phi\mathrm{grad}\mathbf{v} + \mathbf{v} \otimes \mathrm{grad}\phi$
   (ii) $\mathrm{grad}(\mathbf{u} \cdot \mathbf{v}) = (\mathrm{grad}\mathbf{u})^{\mathrm{T}} \mathbf{v} + (\mathrm{grad}\mathbf{v})^{\mathrm{T}} \mathbf{u}$
   (iii) $\mathrm{div}(\mathbf{u} \otimes \mathbf{v}) = (\mathrm{grad}\,\mathbf{u})\mathbf{v} + (\mathrm{div}\,\mathbf{v})\mathbf{u}$
   (iv) $\mathrm{curl}(\mathbf{u} \times \mathbf{v}) = \mathbf{u}\mathrm{div}\mathbf{v} - \mathbf{v}\mathrm{div}\mathbf{u} + (\mathrm{grad}\mathbf{u})\mathbf{v} - (\mathrm{grad}\mathbf{v})\mathbf{u}$
   (v) $\mathrm{div}(\phi\mathbf{A}) = \mathbf{A}\mathrm{grad}\phi + \phi\mathrm{div}\mathbf{A}$
   (vi) $\mathrm{div}(\mathbf{A}\mathbf{v}) = \mathbf{v} \cdot \mathrm{div}\mathbf{A}^{\mathrm{T}} + \mathrm{tr}(\mathbf{A}\mathrm{grad}\mathbf{v})$
   (vii) $\mathrm{div}(\mathbf{AB}) = \mathbf{A}\mathrm{div}\mathbf{B} + \mathrm{grad}\mathbf{A} : \mathbf{B}$
   (viii) $\mathrm{div}(\mathbf{A}(\phi\mathbf{B})) = \phi\mathrm{div}(\mathbf{AB}) + \mathbf{A}(\mathbf{B}\mathrm{grad}\phi)$
   (ix) $\mathrm{grad}(\phi\mathbf{A}) = \phi\mathrm{grad}\mathbf{A} + \mathbf{A} \otimes \mathrm{grad}\phi$
6. Derive Eqn. 1.14.19, the divergence of a tensor in cylindrical coordinates.
7. Deduce the Divergence Theorem identities in 1.14.22 [Hint: write them in index notation.]

1.15 Tensor Calculus 2: Tensor Functions

1.15.1 Vector-valued functions of a vector

Consider a vector-valued function of a vector

$$\mathbf{a} = \mathbf{a}(\mathbf{b}), \qquad a_i = a_i(b_j)$$

This is a function of three independent variables b_1, b_2, b_3, and there are nine partial derivatives $\partial a_i / \partial b_j$. The partial derivative of the vector **a** with respect to **b** is defined to be a second-order tensor with these partial derivatives as its components:

$$\frac{\partial \mathbf{a}(\mathbf{b})}{\partial \mathbf{b}} \equiv \frac{\partial a_i}{\partial b_j} \mathbf{e}_i \otimes \mathbf{e}_j \tag{1.15.1}$$

It follows from this that

$$\frac{\partial \mathbf{a}}{\partial \mathbf{b}} = \left(\frac{\partial \mathbf{b}}{\partial \mathbf{a}} \right)^{-1} \quad \text{or} \quad \frac{\partial \mathbf{a}}{\partial \mathbf{b}} \frac{\partial \mathbf{b}}{\partial \mathbf{a}} = \mathbf{I}, \quad \frac{\partial a_i}{\partial b_m} \frac{\partial b_m}{\partial a_j} = \delta_{ij} \tag{1.15.2}$$

To show this, with $a_i = a_i(b_j),\, b_i = b_i(a_j)$, note that the differential can be written as

$$da_1 = \frac{\partial a_1}{\partial b_j} db_j = \frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_i} da_i = da_1 \left(\frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_1} \right) + da_2 \left(\frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_2} \right) + da_3 \left(\frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_3} \right)$$

Since da_1, da_2, da_3 are independent, one may set $da_2 = da_3 = 0$, so that

$$\frac{\partial a_1}{\partial b_j} \frac{\partial b_j}{\partial a_1} = 1$$

Similarly, the terms inside the other brackets are zero and, in this way, one finds Eqn. 1.15.2.

1.15.2 Scalar-valued functions of a tensor

Consider a scalar valued function of a (second-order) tensor

$$\phi = \phi(\mathbf{T}), \qquad \mathbf{T} = T_{ij} \mathbf{e}_i \otimes \mathbf{e}_j \,.$$

This is a function of nine independent variables, $\phi = \phi(T_{ij})$, so there are nine different partial derivatives:

$$\frac{\partial \phi}{\partial T_{11}}, \frac{\partial \phi}{\partial T_{12}}, \frac{\partial \phi}{\partial T_{13}}, \frac{\partial \phi}{\partial T_{21}}, \frac{\partial \phi}{\partial T_{22}}, \frac{\partial \phi}{\partial T_{23}}, \frac{\partial \phi}{\partial T_{31}}, \frac{\partial \phi}{\partial T_{32}}, \frac{\partial \phi}{\partial T_{33}}$$

The partial derivative of ϕ with respect to $\mathbf{T}$ is defined to be a second-order tensor with these partial derivatives as its components:

$$\boxed{\frac{\partial \phi}{\partial \mathbf{T}} \equiv \frac{\partial \phi}{\partial T_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j} \quad \textbf{Partial Derivative with respect to a Tensor} \qquad (1.15.3)$$

The quantity $\partial \phi(\mathbf{T}) / \partial \mathbf{T}$ is also called the gradient of ϕ with respect to $\mathbf{T}$.

Thus differentiation with respect to a second-order tensor raises the order by 2. This agrees with the idea of the gradient of a scalar field where differentiation with respect to a vector raises the order by 1.

Derivatives of the Trace and Invariants

Consider now the trace: the derivative of $\mathrm{tr}\mathbf{A}$, with respect to $\mathbf{A}$ can be evaluated as follows:

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{A}} \mathrm{tr}\mathbf{A} &= \frac{\partial A_{11}}{\partial \mathbf{A}} + \frac{\partial A_{22}}{\partial \mathbf{A}} + \frac{\partial A_{33}}{\partial \mathbf{A}} \\
&= \frac{\partial A_{11}}{\partial A_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j + \frac{\partial A_{22}}{\partial A_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j + \frac{\partial A_{33}}{\partial A_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j \\
&= \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3 \\
&= \mathbf{I}
\end{aligned} \qquad (1.15.4)$$

Similarly, one finds that {▲Problem 1}

$$\boxed{\begin{array}{ccc}
\dfrac{\partial(\mathrm{tr}\mathbf{A})}{\partial \mathbf{A}} = \mathbf{I} & \dfrac{\partial(\mathrm{tr}\mathbf{A}^2)}{\partial \mathbf{A}} = 2\mathbf{A}^{\mathrm{T}} & \dfrac{\partial(\mathrm{tr}\mathbf{A}^3)}{\partial \mathbf{A}} = 3\left(\mathbf{A}^2\right)^{\mathrm{T}} \\
\dfrac{\partial((\mathrm{tr}\mathbf{A})^2)}{\partial \mathbf{A}} = 2(\mathrm{tr}\mathbf{A})\mathbf{I} & & \dfrac{\partial((\mathrm{tr}\mathbf{A})^3)}{\partial \mathbf{A}} = 3(\mathrm{tr}\mathbf{A})^2 \mathbf{I}
\end{array}} \qquad (1.15.5)$$

Derivatives of Trace Functions

From these and 1.11.17, one can evaluate the derivatives of the invariants {▲Problem 2}:

$$\boxed{\begin{aligned}
\frac{\partial \mathrm{I}_{\mathbf{A}}}{\partial \mathbf{A}} &= \mathbf{I} \\
\frac{\partial \mathrm{II}_{\mathbf{A}}}{\partial \mathbf{A}} &= \mathrm{I}_{\mathbf{A}} \mathbf{I} - \mathbf{A}^{\mathrm{T}} \\
\frac{\partial \mathrm{III}_{\mathbf{A}}}{\partial \mathbf{A}} &= \left(\mathbf{A}^{\mathrm{T}}\right)^2 - \mathrm{I}_{\mathbf{A}} \mathbf{A}^{\mathrm{T}} + \mathrm{II}_{\mathbf{A}} \mathbf{I} = \mathrm{III}_{\mathbf{A}} \mathbf{A}^{-\mathrm{T}}
\end{aligned}} \quad \textbf{Derivatives of the Invariants} \qquad (1.15.6)$$

Derivative of the Determinant

An important relation is

$$\frac{\partial}{\partial \mathbf{A}}(\det \mathbf{A}) = (\det \mathbf{A})\mathbf{A}^{-\mathrm{T}} \tag{1.15.7}$$

which follows directly from 1.15.6c.

Other Relations

The total differential can be written as

$$\begin{aligned} d\phi &= \frac{\partial \phi}{\partial T_{11}} dT_{11} + \frac{\partial \phi}{\partial T_{12}} dT_{12} + \frac{\partial \phi}{\partial T_{13}} dT_{13} + \cdots \\ &\equiv \frac{\partial \phi}{\partial \mathbf{T}} : d\mathbf{T} \end{aligned} \tag{1.15.8}$$

This total differential gives an approximation to the total increment in ϕ when the increments of the independent variables $T_{11}, \cdots$ are small.

The second partial derivative is defined similarly:

$$\frac{\partial^2 \phi}{\partial \mathbf{T} \partial \mathbf{T}} \equiv \frac{\partial^2 \phi}{\partial T_{ij} \partial T_{pq}} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_p \otimes \mathbf{e}_q, \tag{1.15.9}$$

the result being in this case a fourth-order tensor.

Consider a scalar-valued function of a tensor, $\phi(\mathbf{A})$, but now suppose that the components of **A** depend upon some scalar parameter *t*: $\phi = \phi(\mathbf{A}(t))$. By means of the chain rule of differentiation,

$$\dot{\phi} = \frac{\partial \phi}{\partial A_{ij}} \frac{dA_{ij}}{dt} \tag{1.15.10}$$

which in symbolic notation reads (see Eqn. 1.10.10e)

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial \mathbf{A}} : \frac{d\mathbf{A}}{dt} = \mathrm{tr}\left[\left(\frac{\partial \phi}{\partial \mathbf{A}}\right)^{\mathrm{T}} \frac{d\mathbf{A}}{dt}\right] \tag{1.15.11}$$

Identities for Scalar-valued functions of Symmetric Tensor Functions

Let **C** be a symmetric tensor, $\mathbf{C} = \mathbf{C}^{\mathrm{T}}$. Then the partial derivative of $\phi = \phi(\mathbf{C}(\mathbf{T}))$ with respect to **T** can be written as {▲Problem 3}

(1) $\quad \dfrac{\partial \phi}{\partial \mathbf{T}} = 2\mathbf{T}\dfrac{\partial \phi}{\partial \mathbf{C}}$ for $\mathbf{C} = \mathbf{T}^{\mathrm{T}}\mathbf{T}$

(2) $\quad \dfrac{\partial \phi}{\partial \mathbf{T}} = 2\dfrac{\partial \phi}{\partial \mathbf{T}}\mathbf{C}$ for $\mathbf{C} = \mathbf{T}\mathbf{T}^{\mathrm{T}}$ (1.15.12)

(3) $\quad \dfrac{\partial \phi}{\partial \mathbf{T}} = 2\mathbf{T}\dfrac{\partial \phi}{\partial \mathbf{C}} = 2\dfrac{\partial \phi}{\partial \mathbf{C}}\mathbf{T} = \mathbf{T}\dfrac{\partial \phi}{\partial \mathbf{C}} + \dfrac{\partial \phi}{\partial \mathbf{C}}\mathbf{T}$ for $\mathbf{C} = \mathbf{T}\mathbf{T}$ and symmetric $\mathbf{T}$

Scalar-valued functions of a Symmetric Tensor

Consider the expression

$$\mathbf{B} = \frac{\partial \phi(\mathbf{A})}{\partial \mathbf{A}} \qquad B_{ij} = \frac{\partial \phi(A_{ij})}{\partial A_{ij}} \tag{1.15.13}$$

If **A** is a symmetric tensor, there are a number of ways to consider this expression: two possibilities are that ϕ can be considered to be

(i) a symmetric function of the 9 variables A_{ij}

(ii) a function of 6 independent variables: $\phi = \phi(A_{11}, \overline{A}_{12}, \overline{A}_{13}, A_{22}, \overline{A}_{23}, A_{33})$
where

$$\overline{A}_{12} = \frac{1}{2}(A_{12} + A_{21}) = A_{12} = A_{21}$$
$$\overline{A}_{13} = \frac{1}{2}(A_{13} + A_{31}) = A_{13} = A_{31}$$
$$\overline{A}_{23} = \frac{1}{2}(A_{23} + A_{32}) = A_{23} = A_{32}$$

Looking at (i) and writing $\phi = \phi(A_{11}, A_{12}(\overline{A}_{12}), \cdots, A_{21}(\overline{A}_{12}), \cdots)$, one has, for example,

$$\frac{\partial \phi}{\partial \overline{A}_{12}} = \frac{\partial \phi}{\partial A_{12}}\frac{\partial A_{12}}{\partial \overline{A}_{12}} + \frac{\partial \phi}{\partial A_{21}}\frac{\partial A_{21}}{\partial \overline{A}_{12}} = \frac{\partial \phi}{\partial A_{12}} + \frac{\partial \phi}{\partial A_{21}} = 2\frac{\partial \phi}{\partial A_{12}},$$

the last equality following from the fact that ϕ is a symmetrical function of the A_{ij}.

Thus, depending on how the scalar function is presented, one could write

(i) $\quad B_{11} = \dfrac{\partial \phi}{\partial A_{11}}, \quad B_{12} = \dfrac{\partial \phi}{\partial A_{12}}, \quad B_{13} = \dfrac{\partial \phi}{\partial A_{13}}, \quad$ etc.

(ii) $\quad B_{11} = \dfrac{\partial \phi}{\partial A_{11}}, \quad B_{12} = \dfrac{1}{2}\dfrac{\partial \phi}{\partial \overline{A}_{12}}, \quad B_{13} = \dfrac{1}{2}\dfrac{\partial \phi}{\partial \overline{A}_{13}}, \quad$ etc.

1.15.3 Tensor-valued functions of a tensor

The derivative of a (second-order) tensor **A** with respect to another tensor **B** is defined as

$$\frac{\partial \mathbf{A}}{\partial \mathbf{B}} \equiv \frac{\partial A_{ij}}{\partial B_{pq}} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_p \otimes \mathbf{e}_q \tag{1.15.14}$$

and forms therefore a fourth-order tensor. The total differential $d\mathbf{A}$ can in this case be written as

$$d\mathbf{A} = \frac{\partial \mathbf{A}}{\partial \mathbf{B}} : d\mathbf{B} \tag{1.15.15}$$

Consider now

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}} = \frac{\partial A_{ij}}{\partial A_{kl}} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l$$

The components of the tensor are independent, so

$$\frac{\partial A_{11}}{\partial A_{11}} = 1, \quad \frac{\partial A_{11}}{\partial A_{12}} = 0, \quad \cdots \quad \text{etc.} \qquad \boxed{\frac{\partial A_{mn}}{\partial A_{pq}} = \delta_{mp}\delta_{nq}} \tag{1.15.16}$$

and so

$$\frac{\partial \mathbf{A}}{\partial \mathbf{A}} = \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{I}, \tag{1.15.17}$$

the fourth-order identity tensor of Eqn. 1.12.4.

Example

Consider the scalar-valued function ϕ of the tensor **A** and vector **v** (the "dot" can be omitted from the following and similar expressions),

$$\phi(\mathbf{A}, \mathbf{v}) = \mathbf{v} \cdot \mathbf{A}\mathbf{v}$$

The gradient of ϕ with respect to **v** is

$$\frac{\partial \phi}{\partial \mathbf{v}} = \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \cdot \mathbf{A}\mathbf{v} + \mathbf{v} \cdot \mathbf{A}\frac{\partial \mathbf{v}}{\partial \mathbf{v}} = \mathbf{A}\mathbf{v} + \mathbf{v}\mathbf{A} = \left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)\mathbf{v}$$

On the other hand, the gradient of ϕ with respect to **A** is

$$\frac{\partial \phi}{\partial \mathbf{A}} = \mathbf{v} \cdot \frac{\partial \mathbf{A}}{\partial \mathbf{A}} \mathbf{v} = \mathbf{v} \cdot \mathbf{I}\mathbf{v} = \mathbf{v} \otimes \mathbf{v}$$

■

Consider now the derivative of the inverse, $\partial\mathbf{A}^{-1}/\partial\mathbf{A}$. One can differentiate $\mathbf{A}^{-1}\mathbf{A}=\mathbf{0}$ using the product rule to arrive at

$$\frac{\partial\mathbf{A}^{-1}}{\partial\mathbf{A}}\mathbf{A}=-\mathbf{A}^{-1}\frac{\partial\mathbf{A}}{\partial\mathbf{A}}$$

One needs to be careful with derivatives because of the position of the indices in 1.15.14); it looks like a post-operation of both sides with the inverse leads to $\partial\mathbf{A}^{-1}/\partial\mathbf{A}=-\mathbf{A}^{-1}(\partial\mathbf{A}/\partial\mathbf{A})\mathbf{A}^{-1}=-A_{ik}^{-1}A_{jl}^{-1}\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\otimes\mathbf{e}_l$. However, this is not correct (unless **A** is symmetric). Using the index notation (there is no clear symbolic notation), one has

$$\begin{aligned}
&\frac{\partial A_{im}^{-1}}{\partial A_{kl}}A_{mj}=-A_{im}^{-1}\frac{\partial A_{mj}}{\partial A_{kl}} \qquad (\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\otimes\mathbf{e}_l)\\
\to\ &\frac{\partial A_{im}^{-1}}{\partial A_{kl}}A_{mj}A_{jn}^{-1}=-A_{im}^{-1}\frac{\partial A_{mj}}{\partial A_{kl}}A_{jn}^{-1}\\
\to\ &\frac{\partial A_{im}^{-1}}{\partial A_{kl}}\delta_{mn}=-A_{im}^{-1}\delta_{mk}\delta_{jl}A_{jn}^{-1}\\
\to\ &\frac{\partial A_{ij}^{-1}}{\partial A_{kl}}=-A_{ik}^{-1}A_{lj}^{-1} \qquad (\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\otimes\mathbf{e}_l)
\end{aligned} \tag{1.15.18}$$

■

1.15.4 The Directional Derivative

The directional derivative was introduced in §1.6.11. The ideas introduced there can be extended to tensors. For example, the directional derivative of the trace of a tensor **A**, in the direction of a tensor **T**, is

$$\partial_{\mathbf{A}}(\mathrm{tr}\mathbf{A})[\mathbf{T}]=\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\mathrm{tr}(\mathbf{A}+\varepsilon\mathbf{T})=\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}(\mathrm{tr}\mathbf{A}+\varepsilon\mathrm{tr}\mathbf{T})=\mathrm{tr}\mathbf{T} \tag{1.15.19}$$

As a further example, consider the scalar function $\phi(\mathbf{A})=\mathbf{u}\cdot\mathbf{A}\mathbf{v}$, where **u** and **v** are constant vectors. Then

$$\partial_{\mathbf{A}}\phi(\mathbf{A},\mathbf{u},\mathbf{v})[\mathbf{T}]=\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}[\mathbf{u}\cdot(\mathbf{A}+\varepsilon\mathbf{T})\mathbf{v}]=\mathbf{u}\cdot\mathbf{T}\mathbf{v} \tag{1.15.20}$$

Also, the gradient of ϕ with respect to **A** is

$$\frac{\partial\phi}{\partial\mathbf{A}}=\frac{\partial}{\partial\mathbf{A}}(\mathbf{u}\cdot\mathbf{A}\mathbf{v})=\mathbf{u}\otimes\mathbf{v} \tag{1.15.21}$$

and it can be seen that this is an example of the more general relation

$$\partial_{\mathbf{A}} \phi[\mathbf{T}] = \frac{\partial \phi}{\partial \mathbf{A}} : \mathbf{T} \tag{1.15.22}$$

which is analogous to 1.6.41. Indeed,

$$\begin{aligned} \partial_{\mathbf{x}} \phi[\mathbf{w}] &= \frac{\partial \phi}{\partial \mathbf{x}} \cdot \mathbf{w} \\ \partial_{\mathbf{A}} \phi[\mathbf{T}] &= \frac{\partial \phi}{\partial \mathbf{A}} : \mathbf{T} \\ \partial_{\mathbf{u}} \mathbf{v}[\mathbf{w}] &= \frac{\partial \mathbf{v}}{\partial \mathbf{u}} \mathbf{w} \end{aligned} \tag{1.15.23}$$

Example (the Directional Derivative of the Determinant)

It was shown in §1.6.11 that the directional derivative of the determinant of the 2×2 matrix **A**, in the direction of a second matrix **T**, is

$$\partial_{\mathbf{A}} (\det \mathbf{A})[\mathbf{T}] = A_{11}T_{22} + A_{22}T_{11} - A_{12}T_{21} - A_{21}T_{12}$$

This can be seen to be equal to $\det \mathbf{A} \left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T} \right)$, which will now be proved more generally for tensors **A** and **T**:

$$\begin{aligned} \partial_{\mathbf{A}} (\det \mathbf{A})[\mathbf{T}] &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det(\mathbf{A} + \varepsilon \mathbf{T}) \\ &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det\left[\mathbf{A}\left(\mathbf{I} + \varepsilon \mathbf{A}^{-1}\mathbf{T}\right)\right] \\ &= \det \mathbf{A} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \det\left(\mathbf{I} + \varepsilon \mathbf{A}^{-1}\mathbf{T}\right) \end{aligned}$$

The last line here follows from (1.10.16a). Now the characteristic equation for a tensor **B** is given by (1.11.4, 1.11.5),

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) = 0 = \det(\mathbf{B} - \lambda \mathbf{I})$$

where λ_i are the three eigenvalues of **B**. Thus, setting $\lambda = -1$ and $\mathbf{B} = \varepsilon \mathbf{A}^{-1}\mathbf{T}$,

$$\begin{aligned}
\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] &= \det \mathbf{A}\, \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left(1+\lambda_1\big|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\lambda_2\big|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\lambda_3\big|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A}\, \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left(1+\varepsilon\lambda_1\big|_{\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\varepsilon\lambda_2\big|_{\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\varepsilon\lambda_3\big|_{\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A}\, \left(\lambda_1\big|_{\mathbf{A}^{-1}\mathbf{T}} + \lambda_2\big|_{\mathbf{A}^{-1}\mathbf{T}} + \lambda_3\big|_{\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A}\ \mathrm{tr}\left(\mathbf{A}^{-1}\mathbf{T}\right)
\end{aligned}$$

and, from (1.10.10e),

$$\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] = \det \mathbf{A}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right) \tag{1.15.24}$$

■

Example (the Directional Derivative of a vector function)

Consider the n homogeneous algebraic equations $\mathbf{f}(\mathbf{x}) = \mathbf{o}$:

$$\begin{gathered}
f_1(x_1, x_2, \cdots, x_n) = 0 \\
f_2(x_1, x_2, \cdots, x_n) = 0 \\
\cdots \\
f_n(x_1, x_2, \cdots, x_n) = 0
\end{gathered}$$

The directional derivative of $\mathbf{f}$ in the direction of some vector $\mathbf{u}$ is

$$\begin{aligned}
\partial_{\mathbf{x}} \mathbf{f}(\mathbf{x})[\mathbf{u}] &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathbf{f}(\mathbf{z}(\varepsilon)) \qquad (\mathbf{z} = \mathbf{x} + \varepsilon\mathbf{u}) \\
&= \left(\frac{\partial \mathbf{f}(\mathbf{z})}{\partial \mathbf{z}} \frac{d\mathbf{z}}{d\varepsilon}\right)_{\varepsilon=0} \\
&= \mathbf{K}\mathbf{u}
\end{aligned} \tag{1.15.25}$$

where $\mathbf{K}$, called the **tangent matrix** of the system, is

$$\mathbf{K} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & & \partial f_2/\partial x_n \\ \vdots & & & \vdots \\ \partial f_n/\partial x_1 & & \cdots & \partial f_n/\partial x_n \end{bmatrix}, \quad \partial_{\mathbf{x}} \mathbf{f}[\mathbf{u}] = (\mathrm{grad}\mathbf{f})\mathbf{u}$$

which can be compared to (1.15.23c).

■

Properties of the Directional Derivative

The directional derivative is a linear operator and so one can apply the usual product rule. For example, consider the directional derivative of $\mathbf{A}^{-1}$ in the direction of $\mathbf{T}$:

$$\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\left(\mathbf{A}+\varepsilon\mathbf{T}\right)^{-1}$$

To evaluate this, note that $\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\mathbf{A}\right)[\mathbf{T}] = \partial_{\mathbf{A}}\left(\mathbf{I}\right)[\mathbf{T}] = \mathbf{0}$, since **I** is independent of **A**. The product rule then gives $\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}]\mathbf{A} = -\mathbf{A}^{-1}\partial_{\mathbf{A}}\left(\mathbf{A}\right)[\mathbf{T}]$, so that

$$\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}] = -\mathbf{A}^{-1}\partial_{\mathbf{A}}\mathbf{A}[\mathbf{T}]\mathbf{A}^{-1} = -\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1} \tag{1.15.26}$$

Another important property of the directional derivative is the **chain rule**, which can be applied when the function is of the form $\mathbf{f}(\mathbf{x}) = \hat{\mathbf{f}}(\mathbf{B}(\mathbf{x}))$. To derive this rule, consider (see §1.6.11)

$$\mathbf{f}(\mathbf{x}+\mathbf{u}) \approx \mathbf{f}(\mathbf{x}) + \partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}], \tag{1.15.27}$$

where terms of order $o(\mathbf{u})$ have been neglected, i.e.

$$\lim_{|\mathbf{u}|\to 0}\frac{o(\mathbf{u})}{|\mathbf{u}|} = 0.$$

The left-hand side of the previous expression can also be written as

$$\begin{aligned}\hat{\mathbf{f}}(\mathbf{B}(\mathbf{x}+\mathbf{u})) &\approx \hat{\mathbf{f}}(\mathbf{B}(\mathbf{x}) + \partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]) \\ &\approx \hat{\mathbf{f}}(\mathbf{B}(\mathbf{x})) + \partial_{\mathbf{B}}\hat{\mathbf{f}}(\mathbf{B})[\partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]]\end{aligned}$$

Comparing these expressions, one arrives at the chain rule,

$$\boxed{\partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}] = \partial_{\mathbf{B}}\hat{\mathbf{f}}(\mathbf{B})[\partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]]} \qquad \textbf{Chain Rule} \tag{1.15.28}$$

As an application of this rule, consider the directional derivative of $\det\mathbf{A}^{-1}$ in the direction **T**; here, **f** is $\det\mathbf{A}^{-1}$ and $\hat{\mathbf{f}} = \hat{\mathbf{f}}(\mathbf{B}(\mathbf{A}))$. Let $\mathbf{B} = \mathbf{A}^{-1}$ and $\hat{\mathbf{f}} = \det\mathbf{B}$. Then, from Eqns. 1.15.24, 1.15.25, 1.10.3h, f,

$$\begin{aligned}\partial_{\mathbf{A}}\left(\det\mathbf{A}^{-1}\right)[\mathbf{T}] &= \partial_{\mathbf{B}}\left(\det\mathbf{B}\right)[\partial_{\mathbf{A}}\mathbf{A}^{-1}[\mathbf{T}]] \\ &= \left(\det\mathbf{B}\right)\left(\mathbf{B}^{-\mathrm{T}} : \left(-\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}\right)\right) \\ &= -\det\mathbf{A}^{-1}\left(\mathbf{A}^{\mathrm{T}} : \left(\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}\right)\right) \\ &= -\det\mathbf{A}^{-1}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right)\end{aligned} \tag{1.15.29}$$

1.15.5 Formal Treatment of Tensor Calculus

Following on from §1.6.12 and §1.14.6, a scalar function $f : V^2 \to R$ is **differentiable** at $\mathbf{A} \in V^2$ if there exists a second order tensor $Df(\mathbf{A}) \in V^2$ such that

$$f(\mathbf{A}+\mathbf{H}) = f(\mathbf{A}) + Df(\mathbf{A}) : \mathbf{H} + o(\|\mathbf{H}\|) \quad \text{for all} \quad \mathbf{H} \in V^2 \tag{1.15.30}$$

In that case, the tensor $Df(\mathbf{A})$ is called the **derivative** of f at **A**. It follows from this that $Df(\mathbf{A})$ is that tensor for which

$$\partial_{\mathbf{A}} f[\mathbf{B}] = Df(\mathbf{A}) : \mathbf{B} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} f(\mathbf{A}+\varepsilon\mathbf{B}) \quad \text{for all} \quad \mathbf{B} \in V^2 \tag{1.15.31}$$

For example, from 1.15.24,

$$\partial_{\mathbf{A}}(\det\mathbf{A})[\mathbf{T}] = \det\mathbf{A}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right) = \left(\det\mathbf{A}\ \mathbf{A}^{-\mathrm{T}}\right) : \mathbf{T} \tag{1.15.32}$$

from which it follows, from 1.15.31, that

$$\frac{\partial}{\partial\mathbf{A}}\det\mathbf{A} = \det\mathbf{A}\ \mathbf{A}^{-\mathrm{T}} \tag{1.15.33}$$

which is 1.15.7.

Similarly, a tensor-valued function $\mathbf{T} : V^2 \to V^2$ is **differentiable** at $\mathbf{A} \in V^2$ if there exists a fourth order tensor $D\mathbf{T}(\mathbf{A}) \in V^4$ such that

$$\mathbf{T}(\mathbf{A}+\mathbf{H}) = \mathbf{T}(\mathbf{A}) + D\mathbf{T}(\mathbf{A})\mathbf{H} + o(\|\mathbf{H}\|) \quad \text{for all} \quad \mathbf{H} \in V^2 \tag{1.15.34}$$

In that case, the tensor $D\mathbf{T}(\mathbf{A})$ is called the **derivative** of **T** at **A**. It follows from this that $D\mathbf{T}(\mathbf{A})$ is that tensor for which

$$\partial_{\mathbf{A}}\mathbf{T}[\mathbf{B}] = D\mathbf{T}(\mathbf{A}) : \mathbf{B} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{T}(\mathbf{A}+\varepsilon\mathbf{B}) \quad \text{for all} \quad \mathbf{B} \in V^2 \tag{1.15.35}$$

1.15.6 Problems

1. Evaluate the derivatives (use the chain rule for the last two of these)
$$\frac{\partial(\mathrm{tr}\mathbf{A}^2)}{\partial\mathbf{A}},\quad \frac{\partial(\mathrm{tr}\mathbf{A}^3)}{\partial\mathbf{A}},\quad \frac{\partial((\mathrm{tr}\mathbf{A})^2)}{\partial\mathbf{A}},\quad \frac{\partial((\mathrm{tr}\mathbf{A})^2)}{\partial\mathbf{A}}$$
2. Derive the derivatives of the invariants, Eqn. 1.15.5. [Hint: use the Cayley-Hamilton theorem, Eqn. 1.11.15, to express the derivative of the third invariant in terms of the third invariant.]
3. (a) Consider the scalar valued function $\phi = \phi(\mathbf{C}(\mathbf{F}))$, where $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$. Use the chain rule
$$\frac{\partial\phi}{\partial\mathbf{F}} = \frac{\partial\phi}{\partial C_{mn}}\frac{\partial C_{mn}}{\partial F_{ij}}\mathbf{e}_i \otimes \mathbf{e}_j$$
to show that

$$\frac{\partial \phi}{\partial \mathbf{F}} = 2\mathbf{F}\frac{\partial \phi}{\partial \mathbf{C}}, \qquad \frac{\partial \phi}{\partial F_{ij}} = 2F_{ik}\frac{\partial \phi}{\partial C_{kj}}$$

(b) Show also that

$$\frac{\partial \phi}{\partial \mathbf{U}} = 2\mathbf{U}\frac{\partial \phi}{\partial \mathbf{C}} = 2\frac{\partial \phi}{\partial \mathbf{C}}\mathbf{U}$$

for $\mathbf{C} = \mathbf{UU}$ with $\mathbf{U}$ symmetric.

[Hint: for (a), use the index notation: first evaluate $\partial C_{mn} / \partial F_{ij}$ using the product rule, then evaluate $\partial \phi / \partial F_{ij}$ using the fact that **C** is symmetric.]

4. Show that

(a) $\dfrac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} : \mathbf{B} = -\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$, (b) $\dfrac{\partial \mathbf{A}^{-1}}{\partial \mathbf{A}} : \mathbf{A} \otimes \mathbf{A}^{-1} = -\mathbf{A}^{-1} \otimes \mathbf{A}^{-1}$

5. Show that

$$\frac{\partial \mathbf{A}^{\mathrm{T}}}{\partial \mathbf{A}} : \mathbf{B} = \mathbf{B}^{\mathrm{T}}$$

6. By writing the norm of a tensor $|\mathbf{A}|$, 1.10.14, where **A** is symmetric, in terms of the trace (see 1.10.10), show that

$$\frac{\partial |\mathbf{A}|}{\partial \mathbf{A}} = \frac{\mathbf{A}}{|\mathbf{A}|}$$

7. Evaluate

(i) $\partial_{\mathbf{A}}\left(\mathbf{A}^2\right)[\mathbf{T}]$

(ii) $\partial_{\mathbf{A}}\left(\mathrm{tr}\mathbf{A}^2\right)[\mathbf{T}]$ (see 1.10.10e)

8. Derive 1.15.29 by using the definition of the directional derivative and the relation 1.15.7, $\partial(\det \mathbf{A}) / \partial \mathbf{A} = (\det \mathbf{A})\mathbf{A}^{-\mathrm{T}}$.

1.16 Curvilinear Coordinates

1.16.1 The What and Why of Curvilinear Coordinate Systems

Up until now, a rectangular Cartesian coordinate system has been used, and a set of orthogonal unit base vectors $\mathbf{e}_i$ has been employed as the basis for representation of vectors and tensors. More general coordinate systems, called **curvilinear coordinate systems**, can also be used. An example is shown in Fig. 1.16.1: a Cartesian system shown in Fig. 1.16.1a with basis vectors $\mathbf{e}_i$ and a curvilinear system is shown in Fig. 1.16.1b with basis vectors $\mathbf{g}_i$. Some important points are as follows:

1. The Cartesian space can be generated from the coordinate axes x_i; the generated lines (the dotted lines in Fig. 1.16.1) are perpendicular to each other. The base vectors $\mathbf{e}_i$ lie along these lines (they are tangent to them). In a similar way, the curvilinear space is generated from coordinate curves Θ_i; the base vectors $\mathbf{g}_i$ are tangent to these curves.
2. The Cartesian base vectors $\mathbf{e}_i$ are orthogonal to each other and of unit size; in general, the basis vectors $\mathbf{g}_i$ are not orthogonal to each other and are not of unit size.
3. The Cartesian basis is independent of position; the curvilinear basis changes from point to point in the space (the base vectors may change in orientation and/or magnitude).

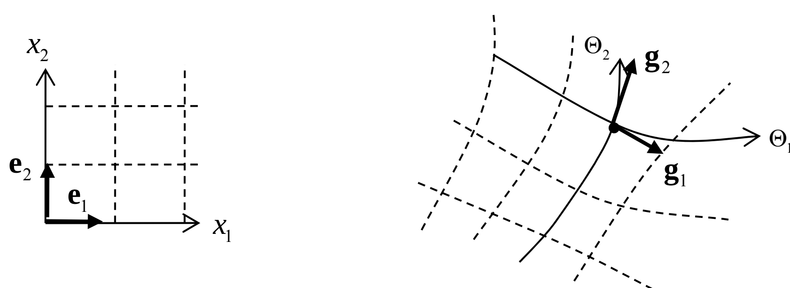


Figure 1.16.1: A Cartesian coordinate system and a curvilinear coordinate system

An example of a curvilinear system is the commonly-used cylindrical coordinate system, shown in Fig. 1.16.2. Here, the curvilinear coordinates $\Theta_1, \Theta_2, \Theta_3$ are the familiar r, θ, z. This cylindrical system is itself a special case of curvilinear coordinates in that the base vectors are always orthogonal to each other. However, unlike the Cartesian system, the orientations of the $\mathbf{g}_1, \mathbf{g}_2$ (r, θ) base vectors change as one moves about the cylinder axis.

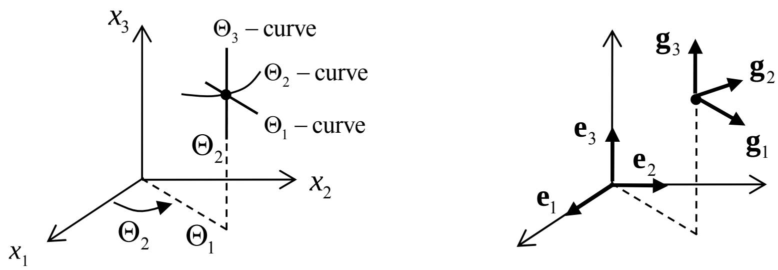


Figure 1.16.2: Cylindrical Coordinates

The question arises: why would one want to use a curvilinear system? There are two main reasons:

1. The problem domain might be of a particular shape, for example a spherical cell, or a soil specimen that is roughly cylindrical. In that case, it is often easier to solve the problems posed by first describing the problem geometry in terms of a curvilinear, e.g. spherical or cylindrical, coordinate system.
2. It may be easier to solve the problem using a Cartesian coordinate system, but a description of the problem in terms of a curvilinear coordinate system allows one to see aspects of the problem which are not obvious in the Cartesian system: it allows for a deeper understanding of the problem.

To give an idea of what is meant by the second point here, consider a simple mechanical deformation of a "square" into a "parallelogram", as shown in Figure 1.16.3. This can be viewed as a deformation of the actual coordinate system, from the Cartesian system aligned with the square, to the "curved" system (actually straight lines, but now not perpendicular) aligned with the edges of the parallelogram. The relationship between the sets of base vectors, the $\mathbf{e}_i$ and the $\mathbf{g}_i$, is intimately connected to the actual physical deformation taking place. In our study of curvilinear coordinates, we will examine this relationship, and also the relationship between the Cartesian coordinates x_i and the curvilinear coordinates Θ_i, and this will give us a very deep knowledge of the essence of the deformation which is taking place. This notion will become more clear when we look at kinematics, convected coordinates and related topics in the next chapter.

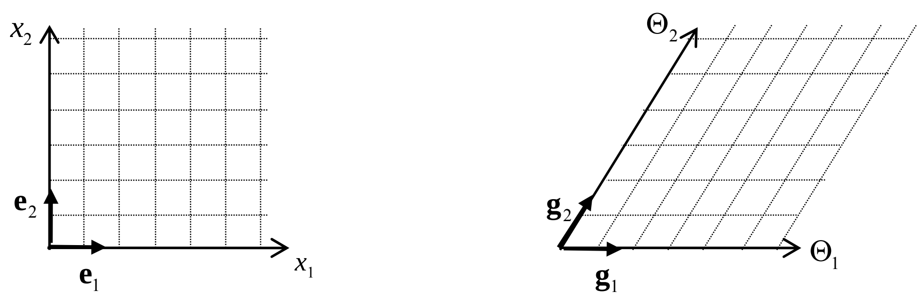


Figure 1.16.3: Deformation of a Square into a Parallelogram

1.16.2 Vectors in Curvilinear Coordinates

The description of scalars in curvilinear coordinates is no different to that in Cartesian coordinates, as they are independent of the basis used. However, the description of vectors is not so straightforward, or obvious, and it will be useful here to work carefully through an example two dimensional problem: consider again the Cartesian coordinate system and the **oblique coordinate system** (which delineates a "parallelogram"-type space), Fig. 1.16.4. These systems have, respectively, base vectors $\mathbf{e}_i$ and $\mathbf{g}_i$, and coordinates x_i and Θ_i. (We will take the $\mathbf{g}_i$ to be of unit size for the purposes of this example.) The base vector $\mathbf{g}_2$ makes an angle α with the horizontal, as shown.

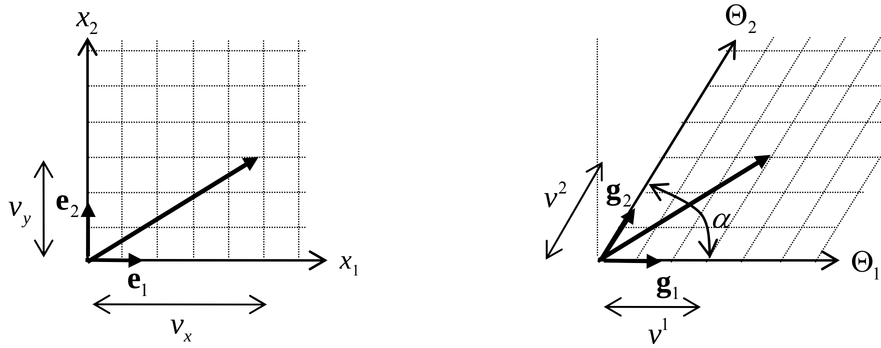


Figure 1.16.4: A Cartesian coordinate system and an oblique coordinate system

Let a vector $\mathbf{v}$ have Cartesian components v_x, v_y, so that it can be described in the Cartesian coordinate system by

$$\mathbf{v} = v_x\mathbf{e}_1 + v_y\mathbf{e}_2 \tag{1.16.1}$$

Let the *same* vector $\mathbf{v}$ have components v^1, v^2 (the reason for using superscripts, when we have always used subscripts hitherto, will become clearer below), so that it can be described in the oblique coordinate system by

$$\mathbf{v} = v^1\mathbf{g}_1 + v^2\mathbf{g}_2 \tag{1.16.2}$$

Using some trigonometry, one can see that these components are related through

$$\begin{aligned} v^1 &= v_x - \frac{1}{\tan\alpha}v_y \\ v^2 &= +\frac{1}{\sin\alpha}v_y \end{aligned} \tag{1.16.3}$$

Now we come to a very important issue: in our work on vector and tensor analysis thus far, a number of important and useful "rules" and relations have been derived. These rules have been *independent of the coordinate system* used. One example is that the magnitude of a vector $\mathbf{v}$ is given by the square root of the dot product: $|\mathbf{v}|^2 = \mathbf{v}\cdot\mathbf{v}$. A natural question to ask is: does this rule work for our oblique coordinate system? To see, first let us evaluate the length squared directly from the Cartesian system:

$$|\mathbf{v}|^2 = \mathbf{v}\cdot\mathbf{v} = v_x^2 + v_y^2 \tag{1.16.4}$$

Following the same logic, we can evaluate

$$\begin{aligned}(v^1)^2 + (v^2)^2 &= \left(v_x - \frac{1}{\tan\alpha}v_y\right)^2 + \left(\frac{1}{\sin\alpha}v_y\right)^2 \\ &= v_x^2 - 2\frac{1}{\tan\alpha}v_x v_y + \left(\frac{1}{\tan^2\alpha} + \frac{1}{\sin^2\alpha}\right)v_y^2\end{aligned} \tag{1.16.5}$$

It is clear from this calculation that our "sum of the squares of the vector components" rule which worked in Cartesian coordinates does not now give us the square of the vector length in curvilinear coordinates.

To find the general rule which works in both (all) coordinate systems, we have to complicate matters somewhat: introduce a *second* set of base vectors into our oblique system. The first set of base vectors, the $\mathbf{g}_1$ and $\mathbf{g}_2$ aligned with the coordinate directions Θ_1 and Θ_2 of Fig. 1.16.4, are termed **covariant** base vectors. Our second set of vectors, which will be termed **contravariant** base vectors, will be denoted by superscripts: $\mathbf{g}^1$ and $\mathbf{g}^2$, and will be aligned with a new set of coordinate directions, Θ^1 and Θ^2.

The new set is defined as follows: the base vector $\mathbf{g}^1$ is perpendicular to $\mathbf{g}_2$ $\left(\mathbf{g}^1\cdot\mathbf{g}_2 = 0\right)$, and the base vector $\mathbf{g}^2$ is perpendicular to $\mathbf{g}_1$ $\left(\mathbf{g}_1\cdot\mathbf{g}^2 = 0\right)$, Fig. 1.16.5a. (The base vectors' orientation with respect to each other follows the "right-hand rule" familiar with Cartesian bases; this will be discussed further below when the general 3D case is examined.) Further, we ensure that

$$\mathbf{g}_1\cdot\mathbf{g}^1 = 1, \qquad \mathbf{g}_2\cdot\mathbf{g}^2 = 1 \tag{1.16.6}$$

With $\mathbf{g}_1 = \mathbf{e}_1$, $\mathbf{g}_2 = \cos\alpha\mathbf{e}_1 + \sin\alpha\mathbf{e}_2$, these conditions lead to

$$\mathbf{g}^1 = \mathbf{e}_1 - \frac{1}{\tan\alpha}\mathbf{e}_2, \quad \mathbf{g}^2 = \frac{1}{\sin\alpha}\mathbf{e}_2 \tag{1.16.7}$$

and $\left|\mathbf{g}^1\right| = \left|\mathbf{g}^2\right| = 1/\sin\alpha$.

A good trick for remembering which are the covariant and which are the contravariant is that the third letter of the word tells us whether the word is associated with subscripts or with superscripts. In "covariant", the "v" is pointing down, so we use subscripts; for "contravariant", the "n" is (with a bit of imagination) pointing up, so we use superscripts.

Let the components of the vector **v** using this new contravariant basis be v_1 and v_2, Fig. 1.16.5b, so that

$$\mathbf{v} = v_1\mathbf{g}^1 + v_2\mathbf{g}^2 \tag{1.16.8}$$

Note the positon of the subscripts and superscripts in this expression: when the base vectors are contravariant ("superscripts"), the associated vector components are covariant ("subscripts"); compare this with the alternative expression for **v** using the covariant basis, Eqn. 1.16.2, $\mathbf{v} = v^1\mathbf{g}_1 + v^2\mathbf{g}_2$, which has covariant base vectors and contravariant vector components.

When **v** is written with covariant components, Eqn. 1.16.8, it is called a **covariant vector**. When **v** is written with contravariant components, Eqn. 1.16.2, it is called a **contravariant vector**. This is not the best of terminology, since it gives the impression that the vector is intrinsically covariant or contravariant, when it is in fact only a matter of which base vectors are being used to describe the vector. For this reason, this terminology will be avoided in what follows.

Examining Fig. 1.16.5b, one can see that $\left|v_1\mathbf{g}^1\right| = v_x / \sin\alpha$ and $\left|v_2\mathbf{g}^2\right| = v_x / \tan\alpha + v_y$, so that

$$\begin{aligned} v_1 &= v_x \\ v_2 &= \cos\alpha\, v_x + \sin\alpha\, v_y \end{aligned} \tag{1.16.9}$$

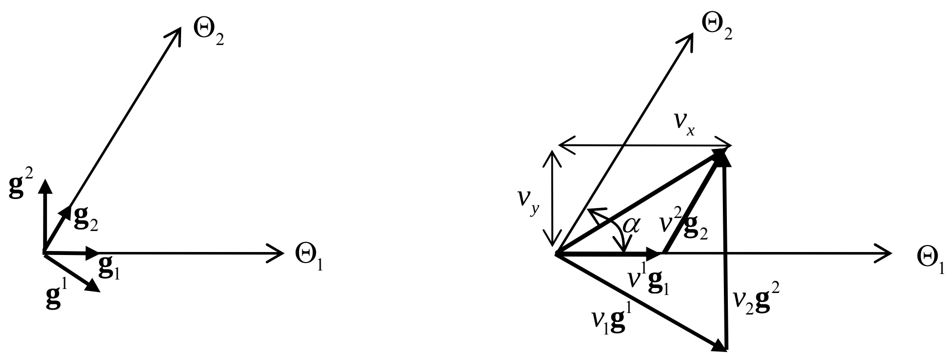


Figure 1.16.5: 2 sets of basis vectors; (a) covariant and contravariant base vectors, (b) covariant and contravariant components of a vector

Now one can evaluate the quantity

$$\begin{aligned} v_1v^1 + v_2v^2 &= \left(v_x\right)\left(v_x - \frac{1}{\tan\alpha}v_y\right) + \left(\cos\alpha\, v_x + \sin\alpha\, v_y\right)\left(\frac{1}{\sin\alpha}v_y\right) \\ &= v_x^2 + v_y^2 \end{aligned} \tag{1.16.10}$$

Thus multiplying the covariant and contravariant components together gives the length squared of the vector; this had to be so given how we earlier defined the two sets of base vectors:

$$
\begin{aligned}
|\mathbf{v}|^2 = \mathbf{v}\cdot\mathbf{v} &= \left(v_1\mathbf{g}^1 + v_2\mathbf{g}^2\right)\cdot\left(v^1\mathbf{g}_1 + v^2\mathbf{g}_2\right) \\
&= v_1v^1\left(\mathbf{g}^1\cdot\mathbf{g}_1\right) + v_2v^2\left(\mathbf{g}^2\cdot\mathbf{g}_2\right) + v_1v^2\left(\mathbf{g}^1\cdot\mathbf{g}_2\right) + v_2v^1\left(\mathbf{g}^2\cdot\mathbf{g}_1\right) \\
&= v_1v^1 + v_2v^2
\end{aligned}
\qquad (1.16.11)
$$

In general, the dot product of two vectors **u** and **v** in the general curvilinear coordinate system is defined through (the fact that the latter equality holds is another consequence of our choice of base vectors, as can be seen by re-doing the calculation of Eqn. 1.16.11 with 2 different vectors, and their different, covariant and contravariant, representations)

$$
\mathbf{u}\cdot\mathbf{v} = u_1v^1 + u_2v^2 = u^1v_1 + u^2v_2 \qquad (1.16.12)
$$

Cartesian Coordinates as Curvilinear Coordinates

The Cartesian coordinate system is a special case of the more general curvilinear coordinate system, where the covariant and contravariant bases are identically the same and the covariant and contravariant components of a vector are identically the same, so that one does not have to bother with carefully keeping track of whether an index is subscript or superscript – we just use subscripts for everything because it is easier.

More formally, in our two-dimensional space, our covariant base vectors are $\mathbf{g}_1 = \mathbf{e}_1, \mathbf{g}_2 = \mathbf{e}_2$. With the contravariant base vectors orthogonal to these, $\mathbf{g}^1\cdot\mathbf{g}_2 = 0$, $\mathbf{g}_1\cdot\mathbf{g}^2 = 0$, and with Eqn. 1.16.6, $\mathbf{g}_1\cdot\mathbf{g}^1 = 1, \mathbf{g}_2\cdot\mathbf{g}^2 = 1$, the contravariant basis is $\mathbf{g}^1 = \mathbf{e}_1, \mathbf{g}^2 = \mathbf{e}_2$. A vector **v** can then be represented as

$$
\mathbf{v} = v_1\mathbf{g}^1 + v_2\mathbf{g}^2 = v^1\mathbf{g}_1 + v^2\mathbf{g}_2 \qquad (1.16.13)
$$

which is nothing other than $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2$, with $v_1 = v^1, v_2 = v^2$. The dot product is, formally,

$$
\mathbf{u}\cdot\mathbf{v} = u_1v^1 + u_2v^2 = u^1v_1 + u^2v_2 \qquad (1.16.14)
$$

which we choose to write as the equivalent $\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2$.

1.16.3 General Curvilinear Coordinates

We now define more generally the concepts discussed above.

A Cartesian coordinate system is defined by the fixed base vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and the coordinates (x^1, x^2, x^3), and any point *p* in space is then determined by the position vector $\mathbf{x} = x^i \mathbf{e}_i$ (see Fig. 1.16.6) [1]. This can be expressed in terms of curvilinear coordinates $(\Theta^1, \Theta^2, \Theta^3)$ by the transformation (and inverse transformation)

$$\begin{aligned} \Theta^i &= \Theta^i\left(x^1, x^2, x^3\right) \\ x^i &= x^i(\Theta^1, \Theta^2, \Theta^3) \end{aligned} \tag{1.16.15}$$

For example, the transformation equations for the oblique coordinate system of Fig. 1.16.4 are

$$\begin{aligned} &\Theta^1 = x^1 - \frac{1}{\tan\alpha} x^2, \quad \Theta^2 = \frac{1}{\sin\alpha} x^2, \quad \Theta^3 = x^3 \\ &x^1 = \Theta^1 + \cos\alpha\,\Theta^2, \quad x^2 = \sin\alpha\,\Theta^2, \quad x^3 = \Theta^3 \end{aligned} \tag{1.16.16}$$

If Θ^1 is varied while holding Θ^2 and Θ^3 constant, a space curve is generated called a Θ^1 **coordinate curve**. Similarly, Θ^2 and Θ^3 coordinate curves may be generated. Three **coordinate surfaces** intersect in pairs along the coordinate curves. On each surface, one of the curvilinear coordinates is constant.

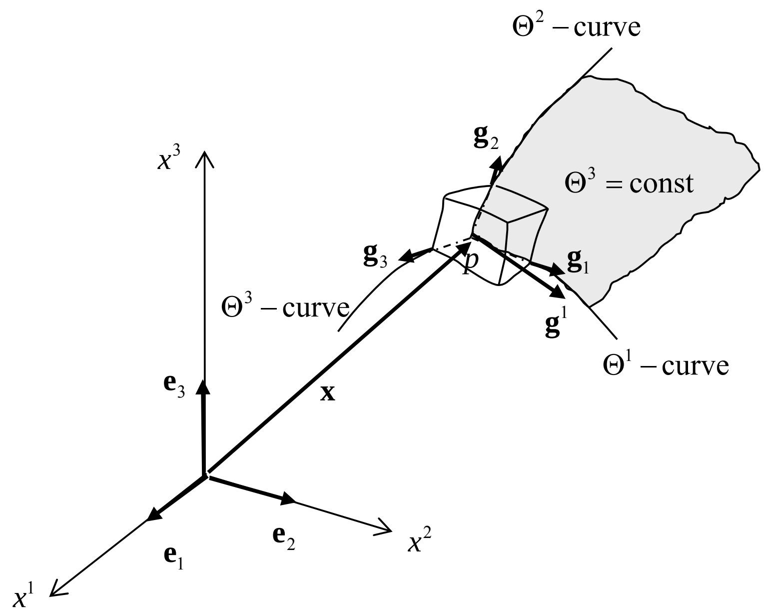


Figure 1.16.6: curvilinear coordinate system and coordinate curves

In order to be able to solve for the Θ^i given the x^i, and to solve for the x^i given the Θ^i, it is necessary and sufficient that the following determinants are non-zero – see Appendix 1.B.2 (the first here is termed the **Jacobian** J of the transformation):

[1] superscripts are used for the Cartesian system here and in much of what follows for notational consistency (see later)

$$J \equiv \det\left[\frac{\partial x^i}{\partial \Theta^j}\right] = \left|\frac{\partial x^i}{\partial \Theta^j}\right|, \qquad \det\left[\frac{\partial \Theta^i}{\partial x^j}\right] = \left|\frac{\partial \Theta^i}{\partial x^j}\right| = \frac{1}{J}, \tag{1.16.17}$$

the last equality following from (1.15.2, 1.10.18d).

Clearly Eqns 1.16.15a can be inverted to get Eqn. 1.16.15b, and vice versa, but just to be sure, we can check that the Jacobian and inverse are non-zero:

$$J = \begin{vmatrix} 1 & \cos\alpha & 0 \\ 0 & \sin\alpha & 0 \\ 0 & 0 & 1 \end{vmatrix} = \sin\alpha, \qquad \frac{1}{J} = \begin{vmatrix} 1 & -\frac{1}{\tan\alpha} & 0 \\ 0 & \frac{1}{\sin\alpha} & 0 \\ 0 & 0 & 1 \end{vmatrix} = \frac{1}{\sin\alpha}, \tag{1.16.18}$$

The Jacobian is zero, i.e. the transformation is singular, only when $\alpha = 0$, i.e. when the parallelogram is shrunk down to a line.

1.16.4 Base Vectors in the Moving Frame

Covariant Base Vectors

From §1.6.2, writing $\mathbf{x} = \mathbf{x}(\Theta^i)$, tangent vectors to the coordinate curves at $\mathbf{x}$ are given by[2]

$$\boxed{\mathbf{g}_i = \frac{\partial \mathbf{x}}{\partial \Theta^i} = \frac{\partial x^m}{\partial \Theta^i}\mathbf{e}_m} \qquad \textbf{Covariant Base Vectors} \tag{1.16.19}$$

with inverse $\mathbf{e}_i = (\partial \Theta^m / \partial x^i)\mathbf{g}_m$. The $\mathbf{g}_i$ emanate from the point *p* and are directed towards the site of increasing coordinate Θ^i. They are called **covariant base vectors**. Increments in the two coordinate systems are related through

$$d\mathbf{x} = \frac{d\mathbf{x}}{\partial \Theta^i} d\Theta^i = \mathbf{g}_i d\Theta^i$$

Note that the triple scalar product $\mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3)$, Eqns. 1.3.17-18, is equivalent to the determinant in 1.16.17,

$$\mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3) = \begin{vmatrix} (\mathbf{g}_1)_1 & (\mathbf{g}_1)_2 & (\mathbf{g}_1)_3 \\ (\mathbf{g}_2)_1 & (\mathbf{g}_2)_2 & (\mathbf{g}_2)_3 \\ (\mathbf{g}_3)_1 & (\mathbf{g}_3)_2 & (\mathbf{g}_3)_3 \end{vmatrix} = J = \det\left[\frac{\partial x^i}{\partial \Theta^j}\right] \tag{1.16.20}$$

[2] in the Cartesian system, with the coordinate curves parallel to the coordinate axes, these equations reduce trivially to $\mathbf{e}_i = (\partial x^m / \partial x^i)\mathbf{e}_m = \delta_{mi}\mathbf{e}_m$

so that the condition that the determinant does not vanish is equivalent to the condition that the vectors $\mathbf{g}_i$ are linearly independent, and so the $\mathbf{g}_i$ can form a basis.

For example, from Eqns. 1.16.16b, the covariant base vectors for the oblique coordinate system of Fig. 1.16.4, are

$$\begin{aligned}
\mathbf{g}_1 &= \frac{\partial x^1}{\partial \Theta^1}\mathbf{e}_1 + \frac{\partial x^2}{\partial \Theta^1}\mathbf{e}_2 + \frac{\partial x^3}{\partial \Theta^1}\mathbf{e}_3 = \mathbf{e}_1 \\
\mathbf{g}_2 &= \frac{\partial x^1}{\partial \Theta^2}\mathbf{e}_1 + \frac{\partial x^2}{\partial \Theta^2}\mathbf{e}_2 + \frac{\partial x^3}{\partial \Theta^2}\mathbf{e}_3 = \cos\alpha\,\mathbf{e}_1 + \sin\alpha\,\mathbf{e}_2 \\
\mathbf{g}_3 &= \frac{\partial x^1}{\partial \Theta^3}\mathbf{e}_1 + \frac{\partial x^2}{\partial \Theta^3}\mathbf{e}_2 + \frac{\partial x^3}{\partial \Theta^3}\mathbf{e}_3 = \mathbf{e}_3
\end{aligned} \tag{1.16.21}$$

Contravariant Base Vectors

Unlike in Cartesian coordinates, where $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, the covariant base vectors do not necessarily form an orthonormal basis, and $\mathbf{g}_i \cdot \mathbf{g}_j \neq \delta_{ij}$. As discussed earlier, in order to deal with this complication, a second set of base vectors are introduced, which are defined as follows: introduce three **contravariant base vectors** $\mathbf{g}^i$ such that each vector is normal to one of the three coordinate surfaces through the point *p*. From §1.6.4, the normal to the coordinate surface $\Theta^1\left(x^1, x^2, x^3\right) = \text{const}$ is given by the gradient vector $\operatorname{grad}\Theta^1$, with Cartesian representation

$$\operatorname{grad}\Theta^1 = \frac{\partial \Theta^1}{\partial x^m}\mathbf{e}^m \tag{1.16.22}$$

and, in general, one may define the contravariant base vectors through

$$\boxed{\mathbf{g}^i = \frac{\partial \Theta^i}{\partial x^m}\mathbf{e}^m} \qquad \textbf{Contravariant Base Vectors} \tag{1.16.23}$$

The contravariant base vector $\mathbf{g}^1$ is shown in Fig. 1.16.6.

As with the covariant base vectors, the triple scalar product $\mathbf{g}^1 \cdot \left(\mathbf{g}^2 \times \mathbf{g}^3\right)$ is equivalent to the determinant in 1.16.17,

$$\mathbf{g}^1 \cdot \left(\mathbf{g}^2 \times \mathbf{g}^3\right) = \begin{vmatrix} \left(\mathbf{g}^1\right)_1 & \left(\mathbf{g}^1\right)_2 & \left(\mathbf{g}^1\right)_3 \\ \left(\mathbf{g}^2\right)_1 & \left(\mathbf{g}^2\right)_2 & \left(\mathbf{g}^2\right)_3 \\ \left(\mathbf{g}^3\right)_1 & \left(\mathbf{g}^3\right)_2 & \left(\mathbf{g}^3\right)_3 \end{vmatrix} = \frac{1}{J} = \det\left[\frac{\partial \Theta^j}{\partial x^i}\right] \tag{1.16.24}$$

and again the condition that the determinant does not vanish is equivalent to the condition that the vectors $\mathbf{g}^i$ are linearly independent, and so the contravariant vectors also form a basis.

From Eqns. 1.16.16a, the contravariant base vectors for the oblique coordinate system are

$$
\begin{aligned}
\mathbf{g}^1 &= \frac{\partial \Theta^1}{\partial x^1}\mathbf{e}_1 + \frac{\partial \Theta^1}{\partial x^2}\mathbf{e}_2 + \frac{\partial \Theta^1}{\partial x^3}\mathbf{e}_3 = \mathbf{e}_1 - \frac{1}{\tan\alpha}\mathbf{e}_2 \\
\mathbf{g}^2 &= \frac{\partial \Theta^2}{\partial x^1}\mathbf{e}_1 + \frac{\partial \Theta^2}{\partial x^2}\mathbf{e}_2 + \frac{\partial \Theta^2}{\partial x^3}\mathbf{e}_3 = \frac{1}{\sin\alpha}\mathbf{e}_2 \\
\mathbf{g}^3 &= \frac{\partial \Theta^3}{\partial x^1}\mathbf{e}_1 + \frac{\partial \Theta^3}{\partial x^2}\mathbf{e}_2 + \frac{\partial \Theta^3}{\partial x^3}\mathbf{e}_3 = \mathbf{e}_3
\end{aligned}
\tag{1.16.25}
$$

1.16.5 Metric Coefficients

It follows from the definitions of the covariant and contravariant vectors that {▲Problem 1}

$$
\boxed{\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j} \tag{1.16.26}
$$

This relation implies that each base vector $\mathbf{g}^i$ is orthogonal to two of the reciprocal base vectors $\mathbf{g}_i$. For example, $\mathbf{g}^1$ is orthogonal to both $\mathbf{g}_2$ and $\mathbf{g}_3$. Eqn. 1.16.26 is the defining relationship between **reciprocal pairs** of general bases. Of course the $\mathbf{g}^i$ were chosen precisely because they satisfy this relation. Here, δ^j_i is again the Kronecker delta[3], with a value of 1 when $i = j$ and zero otherwise.

One needs to be careful to distinguish between subscripts and superscripts when dealing with arbitrary bases, but the rules to follow are straightforward. For example, each free index which is not summed over, such as *i* or *j* in 1.16.26, must be either a subscript or superscript on both sides of an equation. Hence the new notation for the Kronecker delta symbol.

Unlike the orthogonal base vectors, the dot product of a covariant/contravariant base vector with another base vector is not necessarily one or zero. Because of their importance in curvilinear coordinate systems, the dot products are given a special symbol: define the **metric coefficients** to be

$$
\boxed{\begin{aligned} g_{ij} &= \mathbf{g}_i \cdot \mathbf{g}_j \\ g^{ij} &= \mathbf{g}^i \cdot \mathbf{g}^j \end{aligned}} \qquad \textbf{Metric Coefficients} \tag{1.16.27}
$$

For example, the metric coefficients for the oblique coordinate system of Fig. 1.16.4 are

$$
\begin{gathered}
g_{11} = 1, \quad g_{12} = g_{21} = \cos\alpha, \quad g_{22} = 1 \\
g^{11} = \frac{1}{\sin^2\alpha}, \quad g^{12} = g^{21} = -\frac{\cos\alpha}{\sin^2\alpha}, \quad g^{22} = \frac{1}{\sin^2\alpha}
\end{gathered}
\tag{1.16.28}
$$

[3] although in this context it is called the *mixed* Kronecker delta

The following important and useful relations may be derived by manipulating the equations already introduced: {▲Problem 2}

$$\begin{aligned}\mathbf{g}_i &= g_{ij}\mathbf{g}^j \\ \mathbf{g}^i &= g^{ij}\mathbf{g}_j\end{aligned} \tag{1.16.29}$$

and {▲Problem 3}

$$g^{ij}g_{kj} = \delta^i_k \equiv g^i_k \tag{1.16.30}$$

Note here another rule about indices in equations involving general bases: summation can only take place over a dummy index if one is a subscript and the other is a superscript – they are paired off as with the *j*'s in these equations.

The metric coefficients can be written explicitly in terms of the curvilinear components:

$$g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j = \frac{\partial x^k}{\partial \Theta^i}\frac{\partial x^k}{\partial \Theta^j}, \qquad g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j = \frac{\partial \Theta^i}{\partial x^k}\frac{\partial \Theta^j}{\partial x^k} \tag{1.16.31}$$

Note here also a rule regarding derivatives with general bases: the index *i* on the right hand side of 1.16.31a is a superscript of Θ but it is in the denominator of a quotient and so is regarded as a subscript to the entire symbol, matching the subscript *i* on the **g** on the left hand side [4].

One can also write 1.16.31 in the matrix form

$$\left[g_{ij}\right] = \left[\frac{\partial x^k}{\partial \Theta^i}\right]^{\mathrm{T}}\left[\frac{\partial x^k}{\partial \Theta^j}\right], \qquad \left[g^{ij}\right] = \left[\frac{\partial \Theta^i}{\partial x^k}\right]\left[\frac{\partial \Theta^j}{\partial x^k}\right]^{\mathrm{T}} \tag{1.16.32}$$

and, from 1.10.16a,b,

$$\det\left[g_{ij}\right] = \left(\det\left[\frac{\partial x^i}{\partial \Theta^j}\right]\right)^2 = J^2, \qquad \det\left[g^{ij}\right] = \left(\det\left[\frac{\partial \Theta^i}{\partial x^j}\right]\right)^2 = \frac{1}{J^2} \tag{1.16.33}$$

These determinants play an important role, and are denoted by *g*:

$$g = \det\left[g_{ij}\right] = \frac{1}{\det\left[g^{ij}\right]}, \qquad \sqrt{g} = J \tag{1.16.34}$$

<u>Note</u>:

- The matrix $\left[\partial x^k / \partial \Theta^i\right]$ is called the Jacobian *matrix* **J**, so $\mathbf{J}^{\mathrm{T}}\mathbf{J} = \left[g_{ij}\right]$

---

[4] the rule for pairing off indices has been broken in Eqn. 1.16.31 for clarity; more precisely, these equations should be written as $g_{ij} = \left(\partial x^m / \partial \Theta^i\right)\left(\partial x^n / \partial \Theta^j\right)\delta_{mn}$ and $g^{ij} = \left(\partial \Theta^i / \partial x^m\right)\left(\partial \Theta^j / \partial x^n\right)\delta^{mn}$

1.16.6 Scale Factors

The covariant and contravariant base vectors are not necessarily unit vectors. the unit vectors are, with $|\mathbf{g}_i| = \sqrt{\mathbf{g}_i \cdot \mathbf{g}_i}$, $|\mathbf{g}^i| = \sqrt{\mathbf{g}^i \cdot \mathbf{g}^i}$, :

$$\hat{\mathbf{g}}_i = \frac{\mathbf{g}_i}{|\mathbf{g}_i|} = \frac{\mathbf{g}_i}{\sqrt{g_{ii}}}, \qquad \hat{\mathbf{g}}^i = \frac{\mathbf{g}^i}{|\mathbf{g}^i|} = \frac{\mathbf{g}^i}{\sqrt{g^{ii}}} \qquad \text{(no sum)} \tag{1.16.35}$$

The lengths of the covariant base vectors are denoted by *h* and are called the **scale factors**:

$$h_i = |\mathbf{g}_i| = \sqrt{g_{ii}} \qquad \text{(no sum)} \tag{1.16.36}$$

1.16.7 Line Elements and The Metric

Consider a differential line element, Fig. 1.16.7,

$$d\mathbf{x} = dx^i \mathbf{e}_i = d\Theta^i \mathbf{g}_i \tag{1.16.37}$$

The square of the length of this line element, denoted by $(\Delta s)^2$ and called the **metric** of the space, is then

$$(\Delta s)^2 = d\mathbf{x} \cdot d\mathbf{x} = (d\Theta^i \mathbf{g}_i) \cdot (d\Theta^j \mathbf{g}_j) = g_{ij} d\Theta^i d\Theta^j \tag{1.16.38}$$

This relation $(\Delta s)^2 = g_{ij} d\Theta^i d\Theta^j$ is called the **fundamental differential quadratic form**. The g_{ij}'s can be regarded as a set of scale factors for converting increments in Θ^i to changes in length.

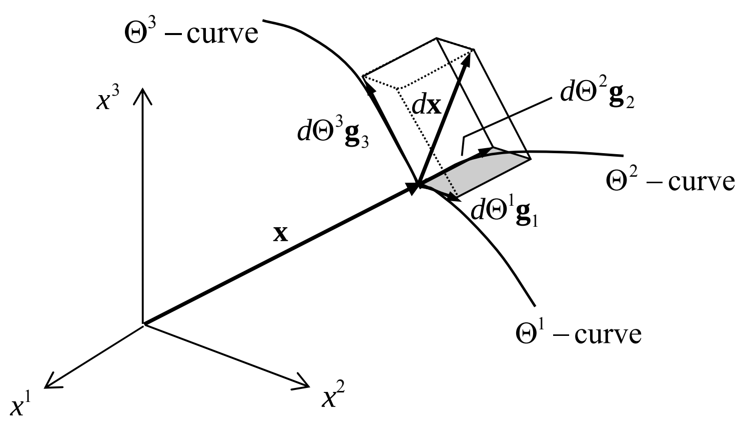


Figure 1.16.7: a line element in space

For a two dimensional space,

$$\begin{aligned}(\Delta s)^2 &= g_{11} d\Theta^1 d\Theta^1 + g_{12} d\Theta^1 d\Theta^2 + g_{21} d\Theta^2 d\Theta^1 + g_{22} d\Theta^2 d\Theta^2 \\ &= g_{11}\left(d\Theta^1\right)^2 + 2g_{12} d\Theta^1 d\Theta^2 + g_{22}\left(d\Theta^2\right)^2\end{aligned} \tag{1.16.39}$$

so that, for the oblique coordinate system of Fig. 1.16.4, from 1.16.28,

$$(\Delta s)^2 = \left(d\Theta^1\right)^2 + 2\cos\alpha\, d\Theta^1 d\Theta^2 + \left(d\Theta^2\right)^2 \tag{1.16.40}$$

This relation can be verified by applying Pythagoras' theorem to the geometry of Figure 1.16.8.

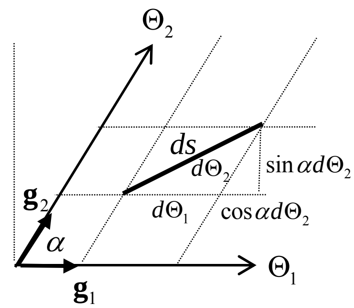


Figure 1.16.8: Length of a line element

1.16.8 Line, Surface and Volume Elements

Here we list expressions for the area of a surface element ΔS and the volume of a volume element ΔV, in terms of the increments in the curvilinear coordinates $\Delta\Theta^1, \Delta\Theta^2, \Delta\Theta^3$. These are particularly useful for the evaluation of surface and volume integrals in curvlinear coordinates.

Surface Area and Volume Elements

The surface area ΔS_1 of a face of the elemental parallelepiped on which Θ_1 is constant (to which $\mathbf{g}^1$ is normal) is, using 1.7.6,

$$\begin{aligned}\Delta S_1 &= \left|(\Delta\Theta^2 \mathbf{g}_2) \times (\Delta\Theta^3 \mathbf{g}_3)\right| \\ &= \left|\mathbf{g}_2 \times \mathbf{g}_3\right| \Delta\Theta^2 \Delta\Theta^3 \\ &= \sqrt{(\mathbf{g}_2 \times \mathbf{g}_3)\cdot(\mathbf{g}_2 \times \mathbf{g}_3)}\,\Delta\Theta^2 \Delta\Theta^3 \\ &= \sqrt{(\mathbf{g}_2\cdot\mathbf{g}_2)(\mathbf{g}_3\cdot\mathbf{g}_3) - (\mathbf{g}_2\cdot\mathbf{g}_3)(\mathbf{g}_2\cdot\mathbf{g}_3)}\,\Delta\Theta^2 \Delta\Theta^3 \\ &= \sqrt{g_{22} g_{33} - (g_{23})^2}\,\Delta\Theta^2 \Delta\Theta^3 \\ &= \sqrt{g g^{11}}\,\Delta\Theta^2 \Delta\Theta^3\end{aligned} \tag{1.16.41}$$

and similarly for the other surfaces. For a two dimensional space, one has

$$\begin{aligned}\Delta S &= \left|(\Delta\Theta^1\mathbf{g}_1)\times(\Delta\Theta^2\mathbf{g}_2)\right| \\ &= \sqrt{g_{11}g_{22}-(g_{12})^2}\,\Delta\Theta^1\Delta\Theta^2 \\ &= \sqrt{g}\,\Delta\Theta^1\Delta\Theta^2 \quad \left(= J\Delta\Theta^1\Delta\Theta^2\right)\end{aligned} \tag{1.16.42}$$

The volume ΔV of the parallelepiped involves the triple scalar product 1.16.20:

$$\Delta V = \left(\mathbf{g}_1\cdot\mathbf{g}_2\times\mathbf{g}_3\right)\Delta\Theta^1\Delta\Theta^2\Delta\Theta^3 = \sqrt{g}\,\Delta\Theta^1\Delta\Theta^2\Delta\Theta^3 \quad \left(= J\Delta\Theta^1\Delta\Theta^2\Delta\Theta^3\right) \tag{1.16.43}$$

1.16.9 Orthogonal Curvilinear Coordinates

In the special case of **orthogonal curvilinear coordinates**, one has

$$g_{ij} = \mathbf{g}_i\cdot\mathbf{g}_j = \delta_{ij}\left|\mathbf{g}_i\right|\left|\mathbf{g}_j\right| = \delta_{ij}h_ih_j, \qquad \left[g_{ij}\right] = \begin{bmatrix} h_1^2 & 0 & 0 \\ 0 & h_2^2 & 0 \\ 0 & 0 & h_3^2\end{bmatrix} \tag{1.16.44}$$

The contravariant base vectors are collinear with the covariant, but the vectors are of different magnitudes:

$$\mathbf{g}_i = h_i\hat{\mathbf{g}}_i, \quad \mathbf{g}^i = \frac{1}{h_i}\hat{\mathbf{g}}_i \tag{1.16.45}$$

It follows that

$$\begin{aligned}(\Delta s)^2 &= h_1^2 d\Theta_1^2 + h_2^2 d\Theta_2^2 + h_3^2 d\Theta_3^2 \\[2ex] \Delta S_1 &= h_2h_3\Delta\Theta_2\Delta\Theta_3 \\ \Delta S_2 &= h_3h_1\Delta\Theta_3\Delta\Theta_1 \\ \Delta S_3 &= h_1h_2\Delta\Theta_1\Delta\Theta_2 \\[2ex] \Delta V &= h_1h_2h_3\Delta\Theta_1\Delta\Theta_2\Delta\Theta_3\end{aligned} \tag{1.16.46}$$

Examples

1. Cylindrical Coordinates

Consider the cylindrical coordinates, $(r,\theta,z) = \left(\Theta^1,\Theta^2,\Theta^3\right)$, *cf.* §1.6.10, Fig. 1.16.9:

$$
\begin{aligned}
x^1 &= \Theta^1 \cos\Theta^2 & \qquad \Theta^1 &= \sqrt{\left(x^1\right)^2 + \left(x^2\right)^2} \\
x^2 &= \Theta^1 \sin\Theta^2 , & \qquad \Theta^2 &= \tan^{-1}\left(x^2 / x^1\right) \\
x^3 &= \Theta^3 & \qquad \Theta^3 &= x^3
\end{aligned}
$$

with

$$
\Theta^1 \geq 0, \quad 0 \leq \Theta^2 < 2\pi, \quad -\infty < \Theta^3 < \infty
$$

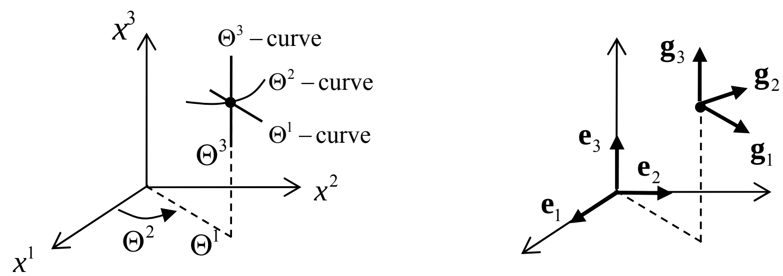


Figure 1.16.9: Cylindrical Coordinates

From Eqns. 1.16.19 (compare with 1.6.29), 1.16.27,

$$
\begin{aligned}
\mathbf{g}_1 &= +\cos\Theta^2 \mathbf{e}_1 + \sin\Theta^2 \mathbf{e}_2 \\
\mathbf{g}_2 &= -\Theta^1 \sin\Theta^2 \mathbf{e}_1 + \Theta^1 \cos\Theta^2 \mathbf{e}_2 , \\
\mathbf{g}_3 &= \mathbf{e}_3
\end{aligned}
\qquad
g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \left(\Theta^1\right)^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

and, from 1.16.17,

$$
J = \det\left[\frac{\partial x^i}{\partial \Theta^j}\right] = \Theta^1
$$

so that there is a one-to-one correspondence between the Cartesian and cylindrical coordinates at all point except for $\Theta^1 = 0$ (which corresponds to the axis of the cylinder). These points are called **singular points** of the transformation. The unit vectors and scale factors are {▲Problem 11}

$$
\begin{aligned}
h_1 &= \left|\mathbf{g}_1\right| = 1 \quad (=1) & \qquad \hat{\mathbf{g}}_1 &= \mathbf{g}_1 \quad (=\mathbf{e}_r) \\
h_2 &= \left|\mathbf{g}_2\right| = \Theta^1 (=r) & \qquad \hat{\mathbf{g}}_2 &= \frac{\mathbf{g}_2}{\Theta^1} (=\mathbf{e}_\theta) \\
h_3 &= \left|\mathbf{g}_3\right| = 1 \quad (=1) & \qquad \hat{\mathbf{g}}_3 &= \mathbf{g}_3 \quad (=\mathbf{e}_z)
\end{aligned}
$$

The line, surface and volume elements are

Metric: $\quad (\Delta s)^2 = \left(d\Theta^1\right)^2 + \left(\Theta^1 d\Theta^2\right)^2 + \left(d\Theta^3\right)^2 \quad \left(= dr^2 + (r d\theta)^2 + dz^2\right)$

$$\Delta S_1 = \Theta^1 \Delta\Theta^2 \Delta\Theta^3$$

Surface Element: $$\Delta S_2 = \quad \Delta\Theta^3 \Delta\Theta^1$$

$$\Delta S_3 = \Theta^1 \Delta\Theta^1 \Delta\Theta^2$$

Volume Element: $$\Delta V = \Theta^1 \Delta\Theta^1 \Delta\Theta^2 \Delta\Theta^3 \quad (= r\Delta r \Delta\theta \Delta z)$$

2. Spherical Coordinates

Consider the spherical coordinates, $(r,\theta,\phi) = (\Theta^1, \Theta^2, \Theta^3)$, *cf.* §1.6.10, Fig. 1.16.10:

$$\begin{aligned} x^1 &= \Theta^1 \sin\Theta^2 \cos\Theta^3 \\ x^2 &= \Theta^1 \sin\Theta^2 \sin\Theta^3 , \\ x^3 &= \Theta^1 \cos\Theta^2 \end{aligned} \qquad \begin{aligned} \Theta_1 &= \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} \\ \Theta_2 &= \tan^{-1}\left(\sqrt{(x^1)^2 + (x^2)^2} / (x^3)^2 \right) \\ \Theta_3 &= \tan^{-1}\left((x^2)^2 / (x^1)^2 \right) \end{aligned}$$

with

$$\Theta_1 \ge 0, \quad 0 \le \Theta^2 \le \pi, \quad 0 \le \Theta^3 < 2\pi$$

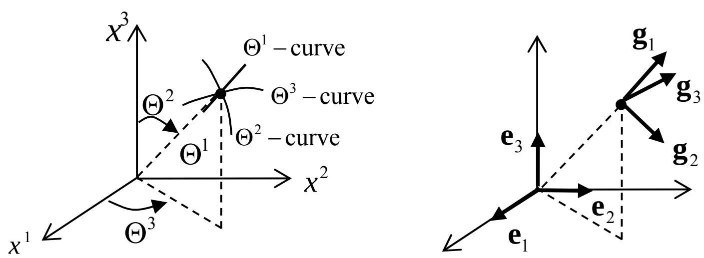


Figure 1.16.10: Spherical Coordinates

From Eqns. 1.16.19 (compare with 1.6.36), 1.16.27,

$$\begin{aligned} \mathbf{g}_1 &= +\sin\Theta^2 \cos\Theta^3 \mathbf{e}_1 + \sin\Theta^2 \sin\Theta^3 \mathbf{e}_2 + \cos\Theta^2 \mathbf{e}_3 \\ \mathbf{g}_2 &= \Theta^1 \left(+\cos\Theta^2 \cos\Theta^3 \mathbf{e}_1 + \cos\Theta^2 \sin\Theta^3 \mathbf{e}_2 - \sin\Theta^2 \mathbf{e}_3 \right), \\ \mathbf{g}_3 &= \Theta^1 \sin\Theta^2 \left(-\sin\Theta^3 \mathbf{e}_1 + \cos\Theta^3 \mathbf{e}_2 \right) \end{aligned} \quad g_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (\Theta^1)^2 & 0 \\ 0 & 0 & (\Theta^1 \sin\Theta^2)^2 \end{bmatrix}$$

and, from 1.16.17,

$$J = \det\left[\frac{\partial x^i}{\partial \Theta^j} \right] = (\Theta^1)^2 \sin\Theta^2$$

so that there is a one-to-one correspondence between the Cartesian and spherical coordinates at all point except for the singular points along the x^3 axis.

The unit vectors and scale factors are {▲Problem 11}

$$\begin{array}{llll} h_1 = |\mathbf{g}_1| = 1 & (=1) & \hat{\mathbf{g}}_1 = \mathbf{g}_1 & (= \mathbf{e}_r) \\ h_2 = |\mathbf{g}_2| = \Theta^1 & (=r) & \hat{\mathbf{g}}_2 = \dfrac{\mathbf{g}_2}{\Theta^1} & (= \mathbf{e}_\theta) \\ h_3 = |\mathbf{g}_3| = \Theta^1 \sin\Theta^2 & (= r\sin\theta) & \hat{\mathbf{g}}_3 = \dfrac{\mathbf{g}_3}{\Theta^1 \sin\Theta^2} & (= \mathbf{e}_\phi) \end{array}$$

The line, surface and volume elements are

Metric:
$$(\Delta s)^2 = (d\Theta^1)^2 + (\Theta^1 d\Theta^2)^2 + (\Theta^1 \sin\Theta^2 d\Theta^3)^2 \quad \left(= dr^2 + (rd\theta)^2 + (r\sin\theta d\phi)^2\right)$$

Surface Element:
$$\begin{aligned} \Delta S_1 &= (\Theta^1)^2 \sin\Theta^2 \Delta\Theta^2 \Delta\Theta^3 \\ \Delta S_2 &= \Theta^1 \sin\Theta^2 \Delta\Theta^3 \Delta\Theta^1 \\ \Delta S_3 &= \Theta^1 \Delta\Theta^1 \Delta\Theta^2 \end{aligned}$$

Volume Element:
$$\Delta V = (\Theta^1)^2 \sin\Theta^2 \Delta\Theta^1 \Delta\Theta^2 \Delta\Theta^3 \quad \left(= r^2 \sin\theta \Delta r \Delta\theta \Delta\phi\right)$$

1.16.10 Vectors in Curvilinear Coordinates

A vector can now be represented in terms of *either* basis:

$$\mathbf{u} = u_i(\Theta^1, \Theta^2, \Theta^3)\mathbf{g}^i = u^i(\Theta^1, \Theta^2, \Theta^3)\mathbf{g}_i \qquad (1.16.47)$$

The u_i are the **covariant components** of **u** and u^i are the **contravariant components** of **u**. Thus the covariant components are the coefficients of the contravariant base vectors and *vice versa* – subscripts denote covariance while superscripts denote contravariance.

Analogous to the orthonormal case, where $\mathbf{u} \cdot \mathbf{e}_i = u_i$ {▲Problem 4}:

$$\mathbf{u} \cdot \mathbf{g}_i = u_i, \quad \mathbf{u} \cdot \mathbf{g}^i = u^i \qquad (1.16.48)$$

Note the following useful formula involving the metric coefficients, for raising or lowering the index on a vector component, relating the covariant and contravariant components, {▲Problem 5}

$$u^i = g^{ij} u_j, \quad u_i = g_{ij} u^j \qquad (1.16.49)$$

Physical Components of a Vector

The contravariant and covariant components of a vector do not have the same physical significance in a curvilinear coordinate system as they do in a rectangular Cartesian system; in fact, they often have different dimensions. For example, the differential $d\mathbf{x}$ of the position vector has in cylindrical coordinates the contravariant components

$(dr, d\theta, dz)$, that is, $d\mathbf{x} = d\Theta^1 \mathbf{g}_1 + d\Theta^2 \mathbf{g}_2 + d\Theta^3 \mathbf{g}_3$ with $\Theta^1 = r$, $\Theta^2 = \theta$, $\Theta^3 = z$. Here, $d\theta$ does not have the same dimensions as the others. The **physical components** in this example are $(dr, rd\theta, dz)$.

The physical components $u^{\langle i \rangle}$ of a vector $\mathbf{u}$ are defined to be the components along the *covariant* base vectors, referred to unit vectors. Thus,

$$\begin{aligned} \mathbf{u} &= u^i \mathbf{g}_i \\ &= \sum_{i=1}^{3} u^i h_i \hat{\mathbf{g}}_i \equiv u^{\langle i \rangle} \hat{\mathbf{g}}_i \end{aligned} \tag{1.16.50}$$

and

$$\boxed{u^{\langle i \rangle} = h_i u^i = \sqrt{g_{ii}}\, u^i} \quad \text{(no sum)} \quad \textbf{Physical Components of a Vector} \tag{1.16.51}$$

From the above, the physical components of a vector $\mathbf{v}$ in the cylindrical coordinate system are $v^1, \Theta^1 v^2, v^3$ and, for the spherical system, $\Theta^1, \Theta^1 v^2, \Theta^1 \sin\Theta^2 v^3$.

The Vector Dot Product

The dot product of two vectors can be written in one of two ways: {▲Problem 6}

$$\boxed{\mathbf{u} \cdot \mathbf{v} = u_i v^i = u^i v_i} \quad \textbf{Dot Product of Two Vectors} \tag{1.16.52}$$

The Vector Cross Product

The triple scalar product is an important quantity in analysis with general bases, particularly when evaluating cross products. From Eqns. 1.16.20, 1.16.24, 1.16.24,

$$\begin{aligned} g &= [\mathbf{g}_1 \cdot \mathbf{g}_2 \times \mathbf{g}_3]^2 = \det[g_{ij}] \\ &= \frac{1}{[\mathbf{g}^1 \cdot \mathbf{g}^2 \times \mathbf{g}^3]^2} = \frac{1}{\det[g^{ij}]} \end{aligned} \tag{1.16.53}$$

Introducing permutation symbols e_{ijk}, e^{ijk}, one can in general write[5]

$$e_{ijk} \equiv \mathbf{g}_i \cdot \mathbf{g}_j \times \mathbf{g}_k = \varepsilon_{ijk} \sqrt{g}, \qquad e^{ijk} \equiv \mathbf{g}^i \cdot \mathbf{g}^j \times \mathbf{g}^k = \varepsilon^{ijk} \frac{1}{\sqrt{g}}$$

---

[5] assuming the base vectors form a *right* handed set, otherwise a negative sign needs to be included

where $\varepsilon_{ijk} = \varepsilon^{ijk}$ is the Cartesian permutation symbol (Eqn. 1.3.10). The cross product of the base vectors can now be written in terms of the reciprocal base vectors as (note the similarity to the Cartesian relation 1.3.12) {▲Problem 7}

$$\boxed{\begin{aligned}\mathbf{g}_i \times \mathbf{g}_j &= e_{ijk}\mathbf{g}^k \\ \mathbf{g}^i \times \mathbf{g}^j &= e^{ijk}\mathbf{g}_k\end{aligned}} \qquad \textbf{Cross Products of Base Vectors} \qquad (1.16.54)$$

Further, from 1.3.19,

$$e^{ijk}e_{pqr} = \varepsilon^{ijk}\varepsilon_{pqr}, \qquad e^{ijk}e_{pqk} = \delta^i_p\delta^j_q - \delta^j_p\delta^i_q \qquad (1.16.55)$$

The Cross Product

The cross product of vectors can be written as {▲Problem 8}

$$\boxed{\begin{aligned}\mathbf{u}\times\mathbf{v} = e_{ijk}u^i v^j \mathbf{g}^k &= \sqrt{g}\begin{vmatrix}\mathbf{g}^1 & \mathbf{g}^2 & \mathbf{g}^3 \\ u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3\end{vmatrix} \\ = e^{ijk}u_i v_j \mathbf{g}_k &= \frac{1}{\sqrt{g}}\begin{vmatrix}\mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3\end{vmatrix}\end{aligned}} \qquad \textbf{Cross Product of Two Vectors} \qquad (1.16.56)$$

1.16.11 Tensors in Curvilinear Coordinates

Tensors can be represented in any of four ways, depending on which combination of base vectors is being utilised:

$$\mathbf{A} = A^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = A_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = A^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j = A_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j \qquad (1.16.56)$$

Here, A^{ij} are the **contravariant components**, A_{ij} are the **covariant components**, $A^i_{\cdot j}$ and $A_i^{\cdot j}$ are the **mixed components** of the tensor **A**. On the mixed components, the subscript is a covariant index, whereas the superscript is called a contravariant index. Note that the "first" index always refers to the first base vector in the tensor product.

An "index switching" rule for tensors is

$$A_{ij}\delta^j_k = A_{ik}, \quad A^{ij}\delta^k_j = A^{ik} \qquad (1.16.57)$$

and the rule for obtaining the components of a tensor **A** is (compare with 1.9.4), {▲Problem 9}

$$\begin{aligned}(\mathbf{A})^{ij} &\equiv A^{ij} = \mathbf{g}^i \cdot \mathbf{A}\mathbf{g}^j \\ (\mathbf{A})_{ij} &\equiv A_{ij} = \mathbf{g}_i \cdot \mathbf{A}\mathbf{g}_j \\ (\mathbf{A})^{i}_{\cdot j} &\equiv A^{i}_{\cdot j} = \mathbf{g}^i \cdot \mathbf{A}\mathbf{g}_j \\ (\mathbf{A})_i^{\cdot j} &\equiv A_i^{\cdot j} = \mathbf{g}_i \cdot \mathbf{A}\mathbf{g}^j\end{aligned} \tag{1.16.58}$$

As with the vectors, the metric coefficients can be used to lower and raise the indices on tensors, for example:

$$\begin{aligned}T^{ij} &= g^{ik} g^{jl} T_{kl} \\ T_i^{\cdot j} &= g_{ik} T^{kj}\end{aligned} \tag{1.16.59}$$

In matrix form, these expressions can be conveniently used to evaluate tensor components, e.g. (note that the matrix of metric coefficients is symmetric)

$$\left[T^{ij}\right] = \left[g^{ik}\right]\left[T_{kl}\right]\left[g^{lj}\right].$$

An example of a higher order tensor is the permutation tensor **E**, whose components are the permutation symbols introduced earlier:

$$\mathbf{E} = e_{ijk} \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k = e^{ijk} \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}_k . \tag{1.16.60}$$

Physical Components of a Tensor

Physical components of tensors can also be defined. For example, if two vectors **a** and **b** have physical components as defined earlier, then the physical components of a tensor **T** are obtained through[6]

$$a^{\langle i \rangle} = T^{\langle ij \rangle} b^{\langle j \rangle} . \tag{1.16.61}$$

As mentioned, physical components are defined with respect to the covariant base vectors, and so the mixed components of a tensor are used, since

$$\mathbf{T}\mathbf{b} = T^{i}_{\cdot j}\left(\mathbf{g}_i \otimes \mathbf{g}^j\right) b^k \mathbf{g}_k = T^{i}_{\cdot j} b^j \mathbf{g}_i \equiv a^i \mathbf{g}_i$$

as required. Then

$$T^{i}_{\cdot j} \frac{b^{\langle j \rangle}}{\sqrt{g_{jj}}} = \frac{a^{\langle i \rangle}}{\sqrt{g_{ii}}} \qquad \text{(no sum on the } g\text{)}$$

and so from 1.16.51,

[6] these are called *right* physical components; *left* physical components are defined through $\mathbf{a} = \mathbf{b}\mathbf{T}$

$$\boxed{T^{\langle ij \rangle} = \frac{\sqrt{g_{ii}}}{\sqrt{g_{jj}}} T^{i}_{\cdot j}} \quad \text{(no sum)} \quad \textbf{Physical Components of a Tensor} \quad (1.16.62)$$

The Identity Tensor

The components of the identity tensor **I** in a general basis can be obtained as follows:

$$\begin{aligned} \mathbf{u} &= u^i \mathbf{g}_i \\ &= g^{ij} u_j \mathbf{g}_i \\ &= g^{ij} (\mathbf{u} \cdot \mathbf{g}_j) \mathbf{g}_i \\ &= g^{ij} (\mathbf{g}_i \otimes \mathbf{g}_j) \mathbf{u} \\ &\equiv \mathbf{Iu} \end{aligned}$$

Thus the contravariant components of the identity tensor are the metric coefficients g^{ij} and, similarly, the covariant components are g_{ij}. For this reason the identity tensor is also called the **metric tensor**. On the other hand, the mixed components are the Kronecker delta, δ^i_j (also denoted by g^i_j). In summary[7],

$$\begin{aligned} (\mathbf{I})_{ij} &= g_{ij} & \mathbf{I} &= g_{ij} \left(\mathbf{g}^i \otimes \mathbf{g}^j \right) \\ (\mathbf{I})^{ij} &= g^{ij} & \mathbf{I} &= g^{ij} \left(\mathbf{g}_i \otimes \mathbf{g}_j \right) \\ (\mathbf{I})^{i}_{\cdot j} &= \delta^i_j & \mathbf{I} &= \delta^i_j \left(\mathbf{g}_i \otimes \mathbf{g}^j \right) = \mathbf{g}_i \otimes \mathbf{g}^i \\ (\mathbf{I})_{i}^{\cdot j} &= \delta_i^j & \mathbf{I} &= \delta_i^j \left(\mathbf{g}^i \otimes \mathbf{g}_j \right) = \mathbf{g}^i \otimes \mathbf{g}_i \end{aligned} \qquad (1.16.63)$$

Symmetric Tensors

A tensor **S** is symmetric if $\mathbf{S}^{\mathrm{T}} = \mathbf{S}$, i.e. if $\mathbf{uSv} = \mathbf{vSu}$. If **S** is symmetric, then

$$S^{ij} = S^{ji}, \quad S_{ij} = S_{ji}, \quad S^{i}_{\cdot j} = S_{j}^{\cdot i} = g_{jk} g^{im} S^{k}_{\cdot m}$$

In terms of matrices,

$$\left[S^{ij} \right] = \left[S^{ij} \right]^T, \quad \left[S_{ij} \right] = \left[S_{ij} \right]^T, \quad \left[S^{i}_{\cdot j} \right] \neq \left[S^{i}_{\cdot j} \right]^{\mathrm{T}}$$

1.16.12 Generalising Cartesian Relations to the Case of General Bases

The tensor relations and definitions already derived for Cartesian vectors and tensors in previous sections, for example in §1.10, are valid also in curvilinear coordinates, for

[7] there is no distinction between δ^j_i, δ^i_j; they are often written as g^j_i, g^i_j and there is no need to specify which index comes first, for example by $g^{\,j}_{\cdot i}$

example $\mathbf{AA}^{-1} = \mathbf{I}$, $\mathrm{tr}\mathbf{A} = \mathbf{I} : \mathbf{A}$ and so on. Formulae involving the index notation may be generalised to arbitrary components by:

(1) raising or lowering the indices appropriately
(2) replacing the (ordinary) Kronecker delta δ_{ij} with the metric coefficients g_{ij}
(3) replacing the Cartesian permutation symbol ε_{ijk} with e_{ijk} in vector cross products

Some examples of this are given in Table 1.16.1 below.

Note that there is only one way of representing a scalar, there are two ways of representing a vector (in terms of its covariant or contravariant components), and there are four ways of representing a (second-order) tensor (in terms of its covariant, contravariant and both types of mixed components).

	Cartesian	General Bases
$\mathbf{a} \cdot \mathbf{b}$	$a_i b_i$	$a_i b^i = a^i b_i$
$\mathbf{aB}$	$a_i B_{ij}$	$(\mathbf{aB})_j = a_i B^i_{\cdot j} = a^i B_{ij}$ $(\mathbf{aB})^j = a_i B^{ij} = a^i B_i^{\cdot j}$
$\mathbf{Ab}$	$A_{ij} b_j$	$(\mathbf{Ab})_i = A_{ij} b^j = A_i^{\cdot j} b_j$ $(\mathbf{Ab})^i = A^{ij} b_j = A^i_{\cdot j} b^j$
$\mathbf{AB}$	$A_{ik} B_{kj}$	$(\mathbf{AB})_{ij} = A_{ik} B^k_{\cdot j} = A_i^{\cdot k} B_{kj}$ $(\mathbf{AB})^{ij} = A^{ik} B_k^{\cdot j} = A^i_{\cdot k} B^{kj}$ $(\mathbf{AB})_i^{\cdot j} = A_i^{\cdot k} B_k^{\cdot j} = A_{ik} B^{kj}$ $(\mathbf{AB})^i_{\cdot j} = A^{ik} B_{kj} = A^i_{\cdot k} B^k_{\cdot j}$
$\mathbf{a} \times \mathbf{b}$	$\varepsilon_{ijk} a_i b_j$	$(\mathbf{a} \times \mathbf{b})_k = e_{ijk} a^i b^j$ $(\mathbf{a} \times \mathbf{b})^k = e^{ijk} a_i b_j$
$\mathbf{a} \otimes \mathbf{b}$	$a_i b_j$	$(\mathbf{a} \otimes \mathbf{b})_{ij} = a_i b_j$ $(\mathbf{a} \otimes \mathbf{b})^{ij} = a^i b^j$ $(\mathbf{a} \otimes \mathbf{b})_i^{\cdot j} = a_i b^j$ $(\mathbf{a} \otimes \mathbf{b})^i_{\cdot j} = a^i b_j$
$\mathbf{A} : \mathbf{B}$	$A_{ij} B_{ij}$	$A_{ij} B^{ij} = A^{ij} B_{ij} = A^i_{\cdot j} B_i^{\cdot j} = A_i^{\cdot j} B^i_{\cdot j}$
$\mathrm{tr}\mathbf{A} \equiv \mathbf{I} : \mathbf{A}$	A_{ii}	$A_i^{\cdot i} = A^i_{\cdot i}$
$\det \mathbf{A}$	$\varepsilon_{ijk} A_{i1} A_{j2} A_{k3}$	$\varepsilon_{ijk} A^i_{\cdot 1} A^j_{\cdot 2} A^k_{\cdot 3}$
$\mathbf{A}^{\mathrm{T}}$	$\left(\mathbf{A}^T\right)_{ij} = A_{ji}$	$\left(\mathbf{A}^{\mathrm{T}}\right)_{ij} = A_{ji}, \left(\mathbf{A}^{\mathrm{T}}\right)^{ij} = A^{ji}$ $\left(\mathbf{A}^{\mathrm{T}}\right)^i_{\cdot j} = A_j^{\cdot i} \neq A^j_{\cdot i}, \left(\mathbf{A}^{\mathrm{T}}\right)_i^{\cdot j} = A^j_{\cdot i} \neq A_j^{\cdot i}$

Table 1.16.1: Tensor relations in Cartesian and general curvilinear coordinates

Rectangular Cartesian (Orthonormal) Coordinate System

In an orthonormal Cartesian coordinate system, $\mathbf{g}_i = \mathbf{g}^i = \mathbf{e}_i$, $g_{ij} = \delta_{ij}$, $g = 1$, $h_i = 1$ and $e_{ijk} = \varepsilon_{ijk}$ $(= \varepsilon^{ijk})$.

1.16.13 Problems

1. Derive the fundamental relation $\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j$.
2. Show that $\mathbf{g}_i = g_{ij}\mathbf{g}^j$ [Hint: assume that one can write $\mathbf{g}_i = a_{ik}\mathbf{g}^k$ and then dot both sides with $\mathbf{g}_j$.]
3. Use the relations 1.16.29 to show that $g^{ij} g_{kj} = \delta^i_k$. Write these equations in matrix form.
4. Show that $\mathbf{u} \cdot \mathbf{g}_i = u_i$.
5. Show that $u_i = g_{ij} u^j$.
6. Show that $\mathbf{u} \cdot \mathbf{v} = u_i v^i = u^i v_i$
7. Use the relation $e_{ijk} \equiv \mathbf{g}_i \cdot \mathbf{g}_j \times \mathbf{g}_k = \varepsilon_{ijk}\sqrt{g}$ to derive the cross product relation $\mathbf{g}_i \times \mathbf{g}_j = e_{ijk}\mathbf{g}^k$. [Hint: show that $\mathbf{g}_i \times \mathbf{g}_j = (\mathbf{g}_i \times \mathbf{g}_j \cdot \mathbf{g}_k)\mathbf{g}^k$.]
8. Derive equation 1.16.56 for the cross product of vectors
9. Show that $(\mathbf{A})_{ij} = \mathbf{g}_i \cdot \mathbf{A}\mathbf{g}_j$.
10. Given $\mathbf{g}_1 = \mathbf{e}_1, \mathbf{g}_2 = \mathbf{e}_2, \mathbf{g}_3 = \mathbf{e}_1 + \mathbf{e}_3$, $\mathbf{v} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$. Find $\mathbf{g}^i, g_{ij}, e_{ijk}, v_i, v^j$ (write the metric coefficients in matrix form).
11. Derive the scale factors for the (a) cylindrical and (b) spherical coordinate systems.
12. **Parabolic Cylindrical** (orthogonal) **coordinates** are given by
$$x^1 = \tfrac{1}{2}\left((\Theta^1)^2 - (\Theta^2)^2\right), \quad x^2 = \Theta^1\Theta^2, \quad x^3 = \Theta^3$$
with
$$-\infty < \Theta^1 < \infty, \quad \Theta^2 \geq 0, \quad -\infty < \Theta^3 < \infty$$
Evaluate:
    (i) the scale factors
    (ii) the Jacobian – are there any singular points?
    (iii) the metric, surface elements, and volume element
Verify that the base vectors $\mathbf{g}_i$ are mutually orthogonal.
[These are intersecting parabolas in the $x^1 - x^2$ plane, all with the same axis]
13. Repeat Problem 7 for the **Elliptical Cylindrical** (orthogonal) **coordinates**:
$$x^1 = a\cosh\Theta^1 \cos\Theta^2, \quad x^2 = a\sinh\Theta^1 \sin\Theta^2, \quad x^3 = \Theta^3$$
with
$$\Theta^1 \geq 0, \quad 0 \leq \Theta^2 < 2\pi, \quad -\infty < \Theta^3 < \infty$$
[These are intersecting ellipses and hyperbolas in the $x^1 - x^2$ plane with foci at $x^1 = \pm a$.]
14. Consider the non-orthogonal curvilinear system illustrated in Fig. 1.16.11, with transformation equations

$$\Theta^1 = x^1 - \frac{1}{\sqrt{3}} x^2$$
$$\Theta^2 = \frac{2}{\sqrt{3}} x^2$$
$$\Theta^3 = x^3$$

Derive the inverse transformation equations, i.e. $x^i = x^i(\Theta^1, \Theta^2, \Theta^3)$, the Jacobian matrices

$$\mathbf{J} = \left[\frac{\partial x^i}{\partial \Theta^j}\right], \quad \mathbf{J}^{-1} = \left[\frac{\partial \Theta^i}{\partial x^j}\right],$$

the covariant and contravariant base vectors, the matrix representation of the metric coefficients $[g_{ij}]$, $[g^{ij}]$, verify that $\mathbf{J}^{\mathrm{T}}\mathbf{J} = [g_{ij}]$, $\mathbf{J}^{-1}\mathbf{J}^{-\mathrm{T}} = [g^{ij}]$ and evaluate g.

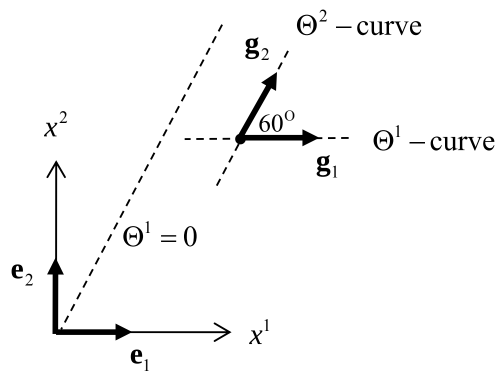


Figure 1.16.11: non-orthogoanl curvilinear coordinate system

15. Consider a (two dimensional) curvilinear coordinate system with covariant base vectors

$$\mathbf{g}_1 = \mathbf{e}_1, \quad \mathbf{g}_2 = \mathbf{e}_1 + \mathbf{e}_2$$

(a) Evaluate the contravariant base vectors and the metric coefficients g_{ij}, g^{ij}

(b) Consider the vectors

$$\mathbf{u} = \mathbf{g}_1 + 3\mathbf{g}_2, \quad \mathbf{v} = -\mathbf{g}_1 + 2\mathbf{g}_2$$

Evaluate the corresponding covariant components of the vectors. Evaluate $\mathbf{u} \cdot \mathbf{v}$ (this can be done in a number of different ways – by using the relations $u_i v^i$, $u^i v_i$, or by directly dotting the vectors in terms of the base vectors $\mathbf{g}_i$, $\mathbf{g}^i$ and using the metric coefficients)

(c) Evaluate the contravariant components of the vector $\mathbf{w} = \mathbf{A}\mathbf{u}$, given that the mixed components $A^i_{\cdot j}$ are

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

Evaluate the contravariant components A^{ij} using the index lowering/raising rule 1.16.59. Re-evaluate the contravariant components of the vector **w** using these components.

16. Consider $\mathbf{A} = A^i_{\cdot j}\mathbf{g}^j \otimes \mathbf{g}_i$. Verify that any of the four versions of **I** in 1.16.63 results in $\mathbf{IA} = \mathbf{I}$.

17. Use the definitions 1.16.19, 1.16.23 to convert $A^{ij}\mathbf{g}_i \otimes \mathbf{g}_j$, $A_{ij}\mathbf{g}^i \otimes \mathbf{g}^j$ and $A^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j$ to the Cartesian bases. Hence show that $\det \mathbf{A}$ is given by the determinant of the matrix of mixed components, $\det\left[A^i_{\cdot j}\right]$, and not by $\det\left[A^{ij}\right]$ or $\det\left[A_{ij}\right]$.

1.17 Curvilinear Coordinates: Transformation Laws

1.17.1 Coordinate Transformation Rules

Suppose that one has a second set of curvilinear coordinates $(\overline{\Theta}^1, \overline{\Theta}^2, \overline{\Theta}^3)$, with

$$\Theta^i = \Theta^i(\overline{\Theta}^1, \overline{\Theta}^2, \overline{\Theta}^3), \qquad \overline{\Theta}^i = \overline{\Theta}^i(\Theta^1, \Theta^2, \Theta^3) \tag{1.17.1}$$

By the chain rule, the covariant base vectors in the second coordinate system are given by

$$\overline{\mathbf{g}}_i = \frac{\partial \mathbf{x}}{\partial \overline{\Theta}^i} = \frac{\partial \Theta^j}{\partial \overline{\Theta}^i}\frac{\partial \mathbf{x}}{\partial \Theta^j} = \frac{\partial \Theta^j}{\partial \overline{\Theta}^i}\mathbf{g}_j$$

A similar calculation can be carried out for the inverse relation and for the contravariant base vectors, giving

$$\begin{aligned} \overline{\mathbf{g}}_i &= \frac{\partial \Theta^j}{\partial \overline{\Theta}^i}\mathbf{g}_j, \qquad & \mathbf{g}_i &= \frac{\partial \overline{\Theta}^j}{\partial \Theta^i}\overline{\mathbf{g}}_j \\ \overline{\mathbf{g}}^i &= \frac{\partial \overline{\Theta}^i}{\partial \Theta^j}\mathbf{g}^j, \qquad & \mathbf{g}^i &= \frac{\partial \Theta^i}{\partial \overline{\Theta}^j}\overline{\mathbf{g}}^j \end{aligned} \tag{1.17.2}$$

The coordinate transformation formulae for vectors $\mathbf{u}$ can be obtained from $\mathbf{u} = u^i\mathbf{g}_i = \overline{u}^i\overline{\mathbf{g}}_i$ and $\mathbf{u} = u_i\mathbf{g}^i = \overline{u}_i\overline{\mathbf{g}}^i$:

$$\boxed{\begin{aligned} \overline{u}^i &= \frac{\partial \overline{\Theta}^i}{\partial \Theta^j}u^j, \qquad & u^i &= \frac{\partial \Theta^i}{\partial \overline{\Theta}^j}\overline{u}^j \\ \overline{u}_i &= \frac{\partial \Theta^j}{\partial \overline{\Theta}^i}u_j, \qquad & u_i &= \frac{\partial \overline{\Theta}^j}{\partial \Theta^i}\overline{u}_j \end{aligned}} \qquad \textbf{Vector Transformation Rule} \tag{1.17.3}$$

These transformation laws have a simple structure and pattern – the subscripts/superscripts on the transformed coordinates $\overline{\Theta}$ quantities match those on the transformed quantities, $\overline{u}, \overline{\mathbf{g}}$, and similarly for the first coordinate system.

<u>Note:</u>

- Covariant and contravariant vectors (and other quantities) are often *defined* in terms of the transformation rules which they obey. For example, a covariant vector can be defined as one whose components transform according to the rules in the second line of the box Eqn. 1.17.3

The transformation laws can be extended to higher-order tensors,

$$\boxed{\begin{aligned}
\overline{A}_{ij} &= \frac{\partial \Theta^m}{\partial \overline{\Theta}^i}\frac{\partial \Theta^n}{\partial \overline{\Theta}^j}A_{mn}, & A_{ij} &= \frac{\partial \overline{\Theta}^m}{\partial \Theta^i}\frac{\partial \overline{\Theta}^n}{\partial \Theta^j}\overline{A}_{mn} \\
\overline{A}^{ij} &= \frac{\partial \overline{\Theta}^i}{\partial \Theta^m}\frac{\partial \overline{\Theta}^j}{\partial \Theta^n}A^{mn}, & A^{ij} &= \frac{\partial \Theta^i}{\partial \overline{\Theta}^m}\frac{\partial \Theta^j}{\partial \overline{\Theta}^n}\overline{A}^{mn} \\
\overline{A}^i_{\cdot j} &= \frac{\partial \overline{\Theta}^i}{\partial \Theta^m}\frac{\partial \Theta^n}{\partial \overline{\Theta}^j}A^m_{\cdot n}, & A^i_{\cdot j} &= \frac{\partial \Theta^i}{\partial \overline{\Theta}^m}\frac{\partial \overline{\Theta}^n}{\partial \Theta^j}\overline{A}^m_{\cdot n} \\
\overline{A}_i^{\cdot j} &= \frac{\partial \overline{\Theta}^j}{\partial \Theta^n}\frac{\partial \Theta^m}{\partial \overline{\Theta}^i}A_m^{\cdot n}, & A_i^{\cdot j} &= \frac{\partial \Theta^j}{\partial \overline{\Theta}^n}\frac{\partial \overline{\Theta}^m}{\partial \Theta^i}\overline{A}_m^{\cdot n}
\end{aligned}}$$

Tensor Transformation Rule (1.17.4)

From these transformation expressions, the following important theorem can be deduced:

If the tensor components are zero in any one coordinate system, they also vanish in any other coordinate system

Reduction to Cartesian Coordinates

For the Cartesian system, let $\mathbf{e}_i = \mathbf{g}_i = \mathbf{g}^i$, $\mathbf{e}'_i = \overline{\mathbf{g}}_i = \overline{\mathbf{g}}^i$ and

$$Q_{ij} = \frac{\partial \Theta^i}{\partial \overline{\Theta}^j} = \frac{\partial x_i}{\partial x'_j} \tag{1.17.5}$$

It follows from 1.17.2 that

$$\frac{\partial \Theta^j}{\partial \overline{\Theta}^i} = \frac{\partial \overline{\Theta}^i}{\partial \Theta^j} \quad \to \quad Q_{ji} = Q^{-1}_{ij} \tag{1.17.6}$$

so the transformation is orthogonal, as expected. Also, as in Eqns. 1.5.11 and 1.5.13.

$$\begin{aligned}
u^i &= \frac{\partial \Theta^i}{\partial \overline{\Theta}^j}\overline{u}^j \quad \to \quad u_i = Q_{ij}u'_j \\
\overline{u}^i &= \frac{\partial \overline{\Theta}^i}{\partial \Theta^j}u^j \quad \to \quad u'_i = Q^{-1}_{ij}u_j = Q_{ji}u_j
\end{aligned} \tag{1.17.7}$$

Transformation Matrix

Transforming coordinates from $\mathbf{g}_i \to \overline{\mathbf{g}}_i$, one can write

$$\mathbf{g}_i = M_i^{\cdot j}\overline{\mathbf{g}}_j = \left(\mathbf{g}_i \cdot \overline{\mathbf{g}}^j\right)\overline{\mathbf{g}}_j \tag{1.17.8}$$

The transformation for a vector can then be expressed, in index notation and matrix notation, as

$$v_i = M_i^{\cdot j}\overline{v}_j, \qquad [v_i] = \left[M_i^{\cdot j}\right]\left[\overline{v}_j\right] \tag{1.17.9}$$

and the transformation matrix is

$$\boxed{\left[M_i^{\cdot j}\right] = \left[\frac{\partial \overline{\Theta}^j}{\partial \Theta^i}\right] = \left[\mathbf{g}_i \cdot \overline{\mathbf{g}}^j\right]} \qquad \textbf{Transformation Matrix} \tag{1.17.10}$$

The rule for contravariant components is then, from 1.17.4,

$$\overline{A}^{ij} = M_m^{\cdot i} M_n^{\cdot j} A^{mn}, \qquad \left[\overline{A}^{ij}\right] = \left[M_m^{\cdot i}\right]^{\mathrm{T}} \left[A^{mn}\right]\left[M_n^{\cdot j}\right] \tag{1.17.11}$$

The Identity Tensor

The identity tensor transforms as

$$\mathbf{I} = \delta_j^i \mathbf{g}_i \otimes \mathbf{g}^j = \mathbf{g}_i \otimes \mathbf{g}^i = \frac{\partial \overline{\Theta}^j}{\partial \Theta^i}\frac{\partial \Theta^i}{\partial \overline{\Theta}^k}\overline{\mathbf{g}}_j \otimes \overline{\mathbf{g}}^k = \delta_k^j \overline{\mathbf{g}}_j \otimes \overline{\mathbf{g}}^k = \overline{\mathbf{g}}_i \otimes \overline{\mathbf{g}}^i \tag{1.17.12}$$

Note that

$$\begin{aligned} g_{ij} &= \mathbf{g}_i \cdot \mathbf{g}_j = \frac{\partial \overline{\Theta}^m}{\partial \Theta^i}\frac{\partial \overline{\Theta}^n}{\partial \Theta^j}\overline{\mathbf{g}}_m \cdot \overline{\mathbf{g}}_n = \frac{\partial \overline{\Theta}^m}{\partial \Theta^i}\frac{\partial \overline{\Theta}^n}{\partial \Theta^j}\overline{g}_{mn} \\ \overline{g}_{ij} &= \overline{\mathbf{g}}_i \cdot \overline{\mathbf{g}}_j = \frac{\partial \Theta^m}{\partial \overline{\Theta}^i}\frac{\partial \Theta^n}{\partial \overline{\Theta}^j}\mathbf{g}_m \cdot \mathbf{g}_n = \frac{\partial \Theta^m}{\partial \overline{\Theta}^i}\frac{\partial \Theta^n}{\partial \overline{\Theta}^j} g_{mn} \end{aligned} \tag{1.17.13}$$

so that, for example,

$$\mathbf{I} = g_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = \frac{\partial \overline{\Theta}^m}{\partial \Theta^i}\frac{\partial \overline{\Theta}^n}{\partial \Theta^j}\overline{g}_{mn}\frac{\partial \Theta^i}{\partial \overline{\Theta}^m}\overline{\mathbf{g}}^m \frac{\partial \Theta^j}{\partial \overline{\Theta}^n} \otimes \overline{\mathbf{g}}^n = \overline{g}_{mn}\overline{\mathbf{g}}^m \otimes \overline{\mathbf{g}}^n \tag{1.17.14}$$

1.17.2 The Metric of the Space

In a second coordinate system, the metric 1.16.38 transforms to

$$\begin{aligned} \overline{(\Delta s)^2} &= \overline{g}_{ij}\overline{\Delta\Theta}^i \overline{\Delta\Theta}^j \\ &= \overline{\mathbf{g}_i \cdot \mathbf{g}_j}\,\overline{\Delta\Theta}^i\overline{\Delta\Theta}^j \\ &= \left(\frac{\partial \Theta^k}{\partial \overline{\Theta}^i}\mathbf{g}_k \cdot \frac{\partial \Theta^m}{\partial \overline{\Theta}^j}\mathbf{g}_m\right)\frac{\partial \overline{\Theta}^i}{\partial \Theta^p}\Delta\Theta^p \frac{\partial \overline{\Theta}^j}{\partial \Theta^q}\Delta\Theta^q \\ &= \delta_p^k \delta_q^m \left(\mathbf{g}_k \cdot \mathbf{g}_m\right)\Delta\Theta^p \Delta\Theta^q \\ &= g_{pq}\Delta\Theta^p\Delta\Theta^q \\ &= \left(\Delta s\right)^2 \end{aligned} \tag{1.17.15}$$

confirming that the metric is a scalar invariant.

1.17.3 Problems

1 Show that $g_{mn}u^m v^n$ is an invariant.

2 How does $\sqrt{g}$ transform between different coordinate systems (in terms of the Jacobian of the transformation, $J = \det\left[\partial\Theta^m / \partial\overline{\Theta}^p\right]$)? [Note that g, although a scalar, is not invariant; it is thus called a **pseudoscalar**.]

3 The components A_{ij} of a tensor **A** are

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

in cylindrical coordinates, at the point $r = 1, \theta = \pi/4, z = \sqrt{3}$. Find the contravariant components of **A** at this point in spherical coordinates. [Hint: use matrix multiplication.]

1.18 Curvilinear Coordinates: Tensor Calculus

1.18.1 Differentiation of the Base Vectors

Differentiation in curvilinear coordinates is more involved than that in Cartesian coordinates because the base vectors are no longer constant and their derivatives need to be taken into account, for example the partial derivative of a vector with respect to the Cartesian coordinates is

$$\frac{\partial \mathbf{v}}{\partial x_j} = \frac{\partial v_i}{\partial x_j}\mathbf{e}_i \qquad \text{but}^1 \qquad \frac{\partial \mathbf{v}}{\partial \Theta^j} = \frac{\partial v^i}{\partial \Theta^j}\mathbf{g}_i + v^i \frac{\partial \mathbf{g}_i}{\partial \Theta^j}$$

The Christoffel Symbols of the Second Kind

First, from Eqn. 1.16.19 – and using the inverse relation,

$$\frac{\partial \mathbf{g}_i}{\partial \Theta^j} = \frac{\partial}{\partial \Theta^j}\left(\frac{\partial x^m}{\partial \Theta^i}\right)\mathbf{e}_m = \frac{\partial^2 x^m}{\partial \Theta^i \Theta^j}\frac{\partial \Theta^k}{\partial x^m}\mathbf{g}_k \tag{1.18.1}$$

this can be written as

$$\boxed{\frac{\partial \mathbf{g}_i}{\partial \Theta^j} = \Gamma_{ij}^k \mathbf{g}_k} \quad \textbf{Partial Derivatives of Covariant Base Vectors} \tag{1.18.2}$$

where

$$\Gamma_{ij}^k = \frac{\partial^2 x^m}{\partial \Theta^i \Theta^j}\frac{\partial \Theta^k}{\partial x^m}, \tag{1.18.3}$$

and Γ_{ij}^k is called the **Christoffel symbol of the second kind**; it can be seen to be equivalent to the *k*th contravariant component of the vector $\partial \mathbf{g}_i / \partial \Theta^j$. One then has {▲Problem 1}

$$\boxed{\Gamma_{ij}^k = \Gamma_{ji}^k = \frac{\partial \mathbf{g}_i}{\partial \Theta^j}\cdot \mathbf{g}^k = \frac{\partial \mathbf{g}_j}{\partial \Theta^i}\cdot \mathbf{g}^k} \quad \textbf{Christoffel Symbols of the 2}^{\textbf{nd}}\textbf{ kind} \tag{1.18.4}$$

and the symmetry in the indices *i* and *j* is evident[2]. Looking now at the derivatives of the contravariant base vectors $\mathbf{g}^i$: differentiating the relation $\mathbf{g}_i \cdot \mathbf{g}^k = \delta_i^k$ leads to

$$-\frac{\partial \mathbf{g}^k}{\partial \Theta^j}\cdot \mathbf{g}_i = \frac{\partial \mathbf{g}_i}{\partial \Theta^j}\cdot \mathbf{g}^k = \Gamma_{ij}^m \mathbf{g}_m \cdot \mathbf{g}^k = \Gamma_{ij}^k$$

---

[1] of course, one could express the $\mathbf{g}_i$ in terms of the $\mathbf{e}_i$, and use only the first of these expressions

[2] note that, in non-Euclidean space, this symmetry in the indices is not necessarily valid

and so

$$\boxed{\frac{\partial \mathbf{g}^i}{\partial \Theta^j} = -\Gamma^i_{jk}\mathbf{g}^k} \quad \textbf{Partial Derivatives of Contravariant Base Vectors} \quad (1.18.5)$$

Transformation formulae for the Christoffel Symbols

The Christoffel symbols are not the components of a (third order) tensor. This follows from the fact that these components do not transform according to the tensor transformation rules given in §1.17. In fact,

$$\overline{\Gamma}^k_{ij} = \frac{\partial \Theta^p}{\partial \overline{\Theta}^i}\frac{\partial \Theta^q}{\partial \overline{\Theta}^j}\frac{\partial \overline{\Theta}^k}{\partial \Theta^r}\Gamma^r_{pq} + \frac{\partial^2 \Theta^s}{\partial \overline{\Theta}^i \overline{\Theta}^j}\frac{\partial \overline{\Theta}^k}{\partial \Theta^s}$$

The Christoffel Symbols of the First Kind

The Christoffel symbols of the second kind relate derivatives of covariant (contravariant) base vectors to the covariant (contravariant) base vectors. A second set of symbols can be introduced relating the base vectors to the derivatives of the reciprocal base vectors, called the **Christoffel symbols of the first kind**:

$$\boxed{\Gamma_{ijk} = \Gamma_{jik} = \frac{\partial \mathbf{g}_i}{\partial \Theta^j}\cdot \mathbf{g}_k = \frac{\partial \mathbf{g}_j}{\partial \Theta^i}\cdot \mathbf{g}_k} \quad \textbf{Christoffel Symbols of the 1}^{\textbf{st}}\textbf{ kind} \quad (1.18.6)$$

so that the partial derivatives of the covariant base vectors can be written in the alternative form

$$\frac{\partial \mathbf{g}_i}{\partial \Theta^j} = \Gamma_{ijk}\mathbf{g}^k, \qquad (1.18.7)$$

and it also follows from Eqn. 1.18.2 that

$$\Gamma_{ijk} = \Gamma^m_{ij} g_{mk}, \quad \Gamma^k_{ij} = \Gamma_{ijm} g^{mk} \qquad (1.18.8)$$

showing that the index k here can be raised or lowered using the metric coefficients as for a third order tensor (but the first two indexes, i and j, cannot and, as stated, the Christoffel symbols are not the components of a third order tensor).

Example: Newton's Second Law

The position vector can be expressed in terms of curvilinear coordinates, $\mathbf{x} = \mathbf{x}(\Theta^i)$. The velocity is then

$$\mathbf{v} = \frac{d\mathbf{x}}{dt} = \frac{\partial \mathbf{x}}{\partial \Theta^i}\frac{d\Theta^i}{dt} = \frac{d\Theta^i}{dt}\mathbf{g}_i$$

and the acceleration is

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\Theta^i}{dt^2}\mathbf{g}_i + \frac{d\Theta^j}{dt}\frac{\partial \mathbf{g}_j}{\partial \Theta^k}\frac{d\Theta^k}{dt} = \left(\frac{d^2\Theta^i}{dt^2} + \Gamma^i_{jk}\frac{d\Theta^j}{dt}\frac{d\Theta^k}{dt}\right)\mathbf{g}_i$$

Equating the contravariant components of Newton's second law $\mathbf{f} = m\mathbf{a}$ then gives the general curvilinear expression

$$f^i = m\left(\ddot{\Theta}^i + \Gamma^i_{jk}\dot{\Theta}^j\dot{\Theta}^k\right)$$

■

Partial Differentiation of the Metric Coefficients

The metric coefficients can be differentiated with the aid of the Christoffel symbols of the first kind {▲Problem 3}:

$$\frac{\partial g_{ij}}{\partial \Theta^k} = \Gamma_{ikj} + \Gamma_{jki} \tag{1.18.9}$$

Using the symmetry of the metric coefficients and the Christoffel symbols, this equation can be written in a number of different ways:

$$g_{ij,k} = \Gamma_{kij} + \Gamma_{jki}, \qquad g_{jk,i} = \Gamma_{ijk} + \Gamma_{kij}, \qquad g_{ki,j} = \Gamma_{jki} + \Gamma_{ijk}$$

Subtracting the first of these from the sum of the second and third then leads to the useful relations (using also 1.18.8)

$$\begin{aligned} \Gamma_{ijk} &= \frac{1}{2}\left(g_{jk,i} + g_{ki,j} - g_{ij,k}\right) \\ \Gamma^k_{ij} &= \frac{1}{2}g^{mk}\left(g_{jm,i} + g_{mi,j} - g_{ij,m}\right) \end{aligned} \tag{1.18.10}$$

which show that the Christoffel symbols depend on the metric coefficients only.

Alternatively, one can write the derivatives of the metric coefficients in the form (the first of these is 1.18.9)

$$\begin{aligned} g_{ij,k} &= \Gamma_{ikj} + \Gamma_{jki} \\ g^{ij}{}_{,k} &= -g^{im}\Gamma^j_{km} - g^{jm}\Gamma^i_{km} \end{aligned} \tag{1.18.11}$$

Also, directly from 1.15.7, one has the relations

$$\frac{\partial g}{\partial g_{ij}} = g g^{ij}, \qquad \frac{\partial g}{\partial g^{ij}} = g g_{ij} \tag{1.18.12}$$

and from these follow other useful relations, for example {▲Problem 4}

$$\Gamma_{ij}^{i} = \frac{\partial \log\left(\sqrt{g}\right)}{\partial \Theta^{j}} = \frac{1}{\sqrt{g}}\frac{\partial \sqrt{g}}{\partial \Theta^{j}} = J^{-1}\frac{\partial J}{\partial \Theta^{j}} \tag{1.18.13}$$

and

$$\begin{aligned} \frac{\partial e_{ijk}}{\partial \Theta^{m}} &= \varepsilon_{ijk}\frac{\partial \sqrt{g}}{\partial \Theta^{m}} = \varepsilon_{ijk}\sqrt{g}\,\Gamma_{mn}^{n} = e_{ijk}\Gamma_{mn}^{n} \\ \frac{\partial e^{ijk}}{\partial \Theta^{m}} &= \varepsilon^{ijk}\frac{\partial\left(1/\sqrt{g}\right)}{\partial \Theta^{m}} = -\varepsilon^{ijk}\frac{1}{\sqrt{g}}\Gamma_{mn}^{n} = -e^{ijk}\Gamma_{mn}^{n} \end{aligned} \tag{1.18.14}$$

1.18.2 Partial Differentiation of Tensors

The Partial Derivative of a Vector

The derivative of a vector in curvilinear coordinates can be written as

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial \Theta^{j}} &= \frac{\partial v^{i}}{\partial \Theta^{j}}\mathbf{g}_{i} + v^{i}\frac{\partial \mathbf{g}_{i}}{\partial \Theta^{j}} & \qquad \frac{\partial \mathbf{v}}{\partial \Theta^{j}} &= \frac{\partial v_{i}}{\partial \Theta^{j}}\mathbf{g}^{i} + v_{i}\frac{\partial \mathbf{g}^{i}}{\partial \Theta^{j}} \\ &= \frac{\partial v^{i}}{\partial \Theta^{j}}\mathbf{g}_{i} + v^{i}\Gamma_{ij}^{k}\mathbf{g}_{k} \qquad \text{or} & &= \frac{\partial v_{i}}{\partial \Theta^{j}}\mathbf{g}^{i} - v_{i}\Gamma_{jk}^{i}\mathbf{g}^{k} \\ &\equiv v^{i}\Big|_{j}\mathbf{g}_{i} & &\equiv v_{i}\Big|_{j}\mathbf{g}^{i} \end{aligned} \tag{1.18.15}$$

where

$$\boxed{\begin{aligned} v^{i}\,|_{j} &= v^{i}{}_{,j} + \Gamma_{kj}^{i}v^{k} \\ v_{i}\,|_{j} &= v_{i,j} - \Gamma_{ij}^{k}v_{k} \end{aligned}} \quad \textbf{Covariant Derivative of Vector Components} \tag{1.18.16}$$

The first term here is the ordinary partial derivative of the vector components. The second term enters the expression due to the fact that the curvilinear base vectors are changing. The complete quantity is defined to be the **covariant derivative** of the vector components. The covariant derivative reduces to the ordinary partial derivative in the case of rectangular Cartesian coordinates.

The $v_i\,|_j$ is the *i*th component of the *j* – derivative of **v**. The $v_i\,|_j$ are also the components of a second order covariant tensor, transforming under a change of coordinate system according to the tensor transformation rule 1.17.4 (see the gradient of a vector below).

The covariant derivative of vector components is given by 1.18.16. In the same way, the covariant derivative of a *vector* is defined to be the complete expression in 1.18.15, $\mathbf{v}_{,j}$, with $\mathbf{v}_{,j} = v^i|_j \, \mathbf{g}_i$.

The Partial Derivative of a Tensor

The rules for covariant differentiation of vectors can be extended to higher order tensors. The various partial derivatives of a second-order tensor

$$\mathbf{A} = A^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = A_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = A_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = A_{\cdot j}^{i}\mathbf{g}_i \otimes \mathbf{g}^j$$

are indicated using the following notation:

$$\frac{\partial \mathbf{A}}{\partial \Theta^k} = A^{ij}|_k \, \mathbf{g}_i \otimes \mathbf{g}_j = A_{ij}|_k \, \mathbf{g}^i \otimes \mathbf{g}^j = A_i^{\cdot j}|_k \, \mathbf{g}^i \otimes \mathbf{g}_j = A_{\cdot j}^{i}|_k \, \mathbf{g}_i \otimes \mathbf{g}^j \qquad (1.18.17)$$

Thus, for example,

$$\begin{aligned}
\mathbf{A}_{,k} &= A_{ij,k}\,\mathbf{g}^i \otimes \mathbf{g}^j + A_{ij}\mathbf{g}^i_{\ ,k} \otimes \mathbf{g}^j + A_{ij}\,\mathbf{g}^i \otimes \mathbf{g}^j_{\ ,k} \\
&= A_{ij,k}\,\mathbf{g}^i \otimes \mathbf{g}^j - A_{ij}\,\Gamma^i_{mk}\mathbf{g}^m \otimes \mathbf{g}^j - A_{ij}\,\mathbf{g}^i \otimes \Gamma^j_{km}\mathbf{g}^m \\
&= \left[A_{ij,k} - \Gamma^m_{ik} A_{mj} - \Gamma^m_{jk} A_{im}\right]\mathbf{g}^i \otimes \mathbf{g}^j
\end{aligned}$$

and, in summary,

$$\boxed{\begin{aligned}
A_{ij}|_k &= A_{ij,k} - \Gamma^m_{ik} A_{mj} - \Gamma^m_{jk} A_{im} \\
A^{ij}|_k &= A^{ij}_{\ ,k} + \Gamma^i_{mk} A^{mj} + \Gamma^j_{mk} A^{im} \\
A_i^{\cdot j}|_k &= A_i^{\cdot j}{}_{,k} + \Gamma^j_{mk} A_i^{\cdot m} - \Gamma^m_{ik} A_m^{\cdot j} \\
A_{\cdot j}^{i}|_k &= A_{\cdot j,k}^{i} + \Gamma^i_{mk} A_{\cdot j}^{m} - \Gamma^m_{jk} A_{\cdot m}^{i}
\end{aligned}} \qquad (1.18.18)$$

Covariant Derivative of Tensor Components

The covariant derivative formulas can be remembered as follows: the formula contains the usual partial derivative plus

- for each contravariant index a term containing a Christoffel symbol in which that index has been inserted on the upper level, multiplied by the tensor component with that index replaced by a dummy summation index which also appears in the Christoffel symbol
- for each covariant index a term prefixed by a minus sign and containing a Christoffel symbol in which that index has been inserted on the lower level, multiplied by the tensor with that index replaced by a dummy which also appears in the Christoffel symbol.
- the remaining symbol in all of the Christoffel symbols is the index of the variable with respect to which the covariant derivative is taken.

For example,

$$A_{\cdot jk}^{i}|_l = A_{\cdot jk,l}^{i} + \Gamma^i_{ml} A_{\cdot jk}^{m} - \Gamma^m_{jl} A_{\cdot mk}^{i} - \Gamma^m_{kl} A_{\cdot jm}^{i}$$

Note that the covariant derivative of a product obeys the same rules as the ordinary differentiation, e.g.

$$\left(u_i A^{jk}\right)|_m = u_i|_m \; A^{jk} + u_i A^{jk}\,|_m$$

Covariantly Constant Coefficients

It can be shown that the metric coefficients are **covariantly constant**[3] {▲Problem 5},

$$g_{ij}|_k = g^{ij}|_k = 0\,,$$

This implies that the metric (identity) tensor **I** is constant, $\mathbf{I}_{,k} = 0$ (see Eqn. 1.16.32) – although its components g_{ij} are not constant. Similarly, the components of the permutation tensor, are covariantly constant

$$e_{ijk}\,|_m = e^{ijk}\,|_m = 0\,.$$

In fact, specialising the identity tensor **I** and the permutation tensor **E** to Cartesian coordinates, one has $g_{ij} = g^{ij} \to \delta_{ij}$, $e_{ijk} = e^{ijk} \to \varepsilon_{ijk}$, which are clearly constant. Specialising the derivatives, $g_{ij}\,|_k \to \delta_{ij,k}$, $e_{ijk}\,|_m \to \varepsilon_{ijk,m}$, and these are clearly zero. From §1.17, since if the components of a tensor vanish in one coordinate system, they vanish in all coordinate systems, the curvilinear coordinate versions vanish also, as stated above.

The above implies that any time any of these factors appears in a covariant derivative, they may be extracted, as in $\left(g_{ij}u^i\right)|_k = \left(g_{ij}\right)u^i\,|_k$.

The Riemann-Christoffel Curvature Tensor

Higher-order covariant derivatives are defined by repeated application of the first-order derivative. This is straight-forward but can lead to algebraically lengthy expressions. For example, to evaluate $v_i|_{mn}$, first write the first covariant derivative in the form of a second order covariant tensor **B**,

$$v_i\,|_m = v_{i,m} - \Gamma^k_{im} v_k \equiv B_{im}$$

so that

$$\begin{aligned} v_i\,|_{mn} &= B_{im}\,|_n \\ &= B_{im,n} - \Gamma^k_{in} B_{km} - \Gamma^k_{mn} B_{ik} \\ &= \left(v_{i,m} - \Gamma^k_{im} v_k\right)_{,n} - \Gamma^k_{in}\left(v_{k,m} - \Gamma^l_{km} v_l\right) - \Gamma^k_{mn}\left(v_{i,k} - \Gamma^l_{ik} v_l\right) \end{aligned} \qquad (1.18.19)$$

3

The covariant derivative $v_i|_{nm}$ is obtained by interchaning *m* and *n* in this expression. Now investigate the difference

$$v_i\,|_{mn} - v_i\,|_{nm} = \left(v_{i,m} - \Gamma^k_{im} v_k\right)_{,n} - \left(v_{i,n} - \Gamma^k_{in} v_k\right)_{,m} - \Gamma^k_{in}\left(v_{k,m} - \Gamma^l_{km} v_l\right) + \Gamma^k_{im}\left(v_{k,n} - \Gamma^l_{kn} v_l\right)$$
$$- \Gamma^k_{mn}\left(v_{i,k} - \Gamma^l_{ik} v_l\right) + \Gamma^k_{nm}\left(v_{i,k} - \Gamma^l_{ik} v_l\right)$$

The last two terms cancel here because of the symmetry of the Christoffel symbol, leaving

$$v_i\,|_{mn} - v_i\,|_{nm} = v_{i,mn} - \Gamma^k_{im,n} v_k - \Gamma^k_{im} v_{k,n} - v_{i,nm} + \Gamma^k_{in,m} v_k + \Gamma^k_{in} v_{k,m}$$
$$- \Gamma^k_{in}\left(v_{k,m} - \Gamma^l_{km} v_l\right) + \Gamma^k_{im}\left(v_{k,n} - \Gamma^l_{kn} v_l\right)$$

The order on the ordinary partial differentiation is interchangeable and so the second order partial derivative terms cancel,

$$v_i\,|_{mn} - v_i\,|_{nm} = v_{i,mn} - \Gamma^k_{im,n} v_k - \Gamma^k_{im} v_{k,n} - v_{i,nm} + \Gamma^k_{in,m} v_k + \Gamma^k_{in} v_{k,m}$$
$$- \Gamma^k_{in}\left(v_{k,m} - \Gamma^l_{km} v_l\right) + \Gamma^k_{im}\left(v_{k,n} - \Gamma^l_{kn} v_l\right)$$

After further cancellation one arrives at

$$v_i|_{mn} - v_i|_{nm} = R^{\cdot j}_{i\,mn} v_j \qquad (1.18.20)$$

where **R** is the fourth-order **Riemann-Christoffel curvature tensor**, with (mixed) components

$$R^{\cdot j}_{i\,mn} = \Gamma^j_{in,m} - \Gamma^j_{im,n} + \Gamma^k_{in}\Gamma^j_{km} - \Gamma^k_{im}\Gamma^j_{kn} \qquad (1.18.21)$$

Since the Christoffel symbols vanish in a Cartesian coordinate system, then so does $R^{\cdot j}_{i\,mn}$. Again, any tensor that vanishes in one coordinate system must be zero in all coordinate systems, and so $R^{\cdot j}_{i\,mn} = 0$, implying that the order of covariant differentiation is immaterial, $v_i|_{mn} = v_i|_{nm}$.

From 1.18.10, it follows that

$$R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij}$$
$$R_{ijkl} + R_{iklj} + R_{iljk} = 0$$

The latter of these known as the **Bianchi identities**. In fact, only six components of the Riemann-Christoffel tensor are independent; the expression $R^{\cdot j}_{i\,mn} = 0$ then represents 6 equations in the 6 independent components g_{ij}.

This analysis is for a Euclidean space – the usual three-dimensional space in which quantities can be expressed in terms of a Cartesian reference system – such a space is called a **flat space**. These ideas can be extended to other, curved spaces, so-called **Riemannian spaces** (**Riemannian manifolds**), for which the Riemann-Christoffel tensor is non-zero (see §1.19).

1.18.3 Differential Operators and Tensors

In this section, the concepts of the gradient, divergence and curl from §1.6 and §1.14 are generalized to the case of curvilinear components.

Space Curves and the Gradient

Consider first a scalar function $f(\mathbf{x})$, where $\mathbf{x} = x^i \mathbf{e}_i$ is the position vector, with $x^i = x^i(\Theta^j)$. Let the curvilinear coordinates depend on some parameter s, $\Theta^j = \Theta^j(s)$, so that $\mathbf{x}(s)$ traces out a space curve C.

For example, the cylindrical coordinates $\Theta^j = \Theta^j(s)$, with $r = a$, $\theta = s/c$, $z = sb/c$, $0 \le s \le 2\pi c$, generate a helix.

From §1.6.2, a tangent to C is

$$\boldsymbol{\tau} = \frac{d\mathbf{x}}{ds} = \frac{\partial \mathbf{x}}{\partial \Theta^i}\frac{d\Theta^i}{ds} = \tau^i \mathbf{g}_i$$

so that $d\Theta^i / ds$ are the contravariant components of $\boldsymbol{\tau}$. Thus

$$\frac{df}{ds} = \frac{\partial f}{\partial \Theta^i}\tau^i = \left(\frac{\partial f}{\partial \Theta^i}\mathbf{g}^i\right)\cdot\left(\tau^j \mathbf{g}_j\right).$$

For Cartesian coordinates, $df/ds = \nabla f \cdot \boldsymbol{\tau}$ (see the discussion on normals to surfaces in §1.6.4). For curvilinear coordinates, therefore, the Nabla operator of 1.6.11 now reads

$$\nabla = \mathbf{g}^i \frac{\partial}{\partial \Theta^i} \tag{1.18.22}$$

so that again the directional derivative is

$$\frac{df}{ds} = \nabla f \cdot \boldsymbol{\tau}$$

The Gradient of a Scalar

In general then, the gradient of a scalar valued function Φ is defined to be

$$\boxed{\nabla\Phi \equiv \mathrm{grad}\Phi = \frac{\partial \Phi}{\partial \Theta^i}\mathbf{g}^i} \quad \textbf{Gradient of a Scalar} \tag{1.18.23}$$

and, with $d\mathbf{x} = dx^i \mathbf{e}_i = d\Theta^i \mathbf{g}_i$, one has

$$d\Phi \equiv \frac{\partial \Phi}{\partial \Theta^i} d\Theta^i = \nabla\Phi \cdot d\mathbf{x} \qquad (1.18.24)$$

The Gradient of a Vector

Analogous to Eqn. 1.14.3, the gradient of a vector is defined to be the tensor product of the derivative $\partial \mathbf{u} / \Theta^j$ with the *contravariant* base vector $\mathbf{g}^j$:

$$\boxed{\begin{aligned}\mathrm{grad}\mathbf{u} = \frac{\partial \mathbf{u}}{\partial \Theta^j} \otimes \mathbf{g}^j &= u_i \,|_j\, \mathbf{g}^i \otimes \mathbf{g}^j \\ &= u^i \,|_j\, \mathbf{g}_i \otimes \mathbf{g}^j\end{aligned}} \quad \textbf{Gradient of a Vector} \qquad (1.18.25)$$

Note that

$$\nabla \otimes \mathbf{u} = \mathbf{g}^i \frac{\partial}{\partial \Theta^i} \otimes \mathbf{u} = \mathbf{g}^i \otimes \frac{\partial \mathbf{u}}{\partial \Theta^i} = u_j \,|_i\, \mathbf{g}^i \otimes \mathbf{g}^j = u^j \,|_i\, \mathbf{g}^i \otimes \mathbf{g}_j$$

so that again one arrives at Eqn. 1.14.7, $(\nabla \otimes \mathbf{u})^{\mathrm{T}} = \mathrm{grad}\mathbf{u}$.

Again, one has for a space curve parameterised by s,

$$\frac{d\mathbf{u}}{ds} = \frac{\partial \mathbf{u}}{\partial \Theta^i} \tau^i = \frac{\partial \mathbf{u}}{\partial \Theta^i} (\boldsymbol{\tau} \cdot \mathbf{g}^i) = \left(\mathbf{g}^i \otimes \frac{\partial \mathbf{u}}{\partial \Theta^i} \right)^{\mathrm{T}} \cdot \boldsymbol{\tau} = \mathrm{grad}\mathbf{u} \cdot \boldsymbol{\tau}$$

Similarly, from 1.18.18, the gradient of a second-order tensor is

$$\boxed{\begin{aligned}\mathrm{grad}\mathbf{A} = \frac{\partial \mathbf{A}}{\partial \Theta^k} \otimes \mathbf{g}^k &= A^{ij} \,|_k\, \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \\ &= A_{ij} \,|_k\, \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \\ &= A_i^{\cdot j} \,|_k\, \mathbf{g}^i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \\ &= A^i_{\cdot j} \,|_k\, \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k\end{aligned}} \quad \textbf{Gradient of a Tensor} \qquad (1.18.26)$$

The Divergence

From 1.14.9, the divergence of a vector is {▲Problem 6}

$$\boxed{\mathrm{div}\mathbf{u} = \mathrm{grad}\,\mathbf{u} : \mathbf{I} = u^i \,|_i \qquad \left(= \frac{\partial \mathbf{u}}{\partial \Theta^j} \cdot \mathbf{g}^j \right)} \quad \textbf{Divergence of a Vector} \qquad (1.18.27)$$

This is equivalent to the divergence operation involving the Nabla operator, $\mathrm{div}\mathbf{u} = \nabla \cdot \mathbf{u}$. An alternative expression can be obtained from 1.18.13 {▲Problem 7},

$$\mathrm{div}\mathbf{u} = u^i\,|_i = \frac{1}{\sqrt{g}}\frac{\partial\left(\sqrt{g}u^i\right)}{\partial\Theta^i} = J^{-1}\frac{\partial\left(Ju^i\right)}{\partial\Theta^i}$$

Similarly, using 1.14.12, the divergence of a second-order tensor is

$$\boxed{\begin{aligned}\mathrm{div}\mathbf{A} &= \mathrm{grad}\,\mathbf{A} : \mathbf{I} = A^{ij}\,|_j\;\mathbf{g}_i \qquad \left(= \frac{\partial\mathbf{A}}{\partial\Theta^j}\mathbf{g}^j\right)\\ &= A_i^{\cdot j}\,|_j\;\mathbf{g}^i\end{aligned}}$$

Divergence of a Tensor (1.18.28)

Here, one has the alternative definition,

$$\nabla\cdot\mathbf{A} = \mathbf{g}^i\frac{\partial}{\partial\Theta^i}\cdot\mathbf{A} = \mathbf{g}^i\cdot\frac{\partial\mathbf{A}}{\partial\Theta^i} = A^{ji}\,|_j\;\mathbf{g}_i = \cdots$$

so that again one arrives at Eqn. 1.14.14, $\mathrm{div}\mathbf{A} = \nabla\cdot\mathbf{A}^{\mathrm{T}}$.

The Curl

The curl of a vector is defined by {▲Problem 8}

$$\boxed{\mathrm{curl}\mathbf{u} = \nabla\times\mathbf{u} = \mathbf{g}^k\times\frac{\partial\mathbf{u}}{\partial\Theta^k} = e^{ijk}u_j\,|_i\;\mathbf{g}_k = e^{ijk}\frac{\partial u_j}{\partial\Theta^i}\mathbf{g}_k}$$

Curl of a Vector (1.18.29)

the last equality following from the fact that all the Christoffel symbols cancel out.

Covariant derivatives as Tensor Components

Equation 1.18.25 shows clearly that the covariant derivatives of vector components are themselves the components of second order tensors. It follows that they can be manipulated as other tensors, for example,

$$g^{im}u_m\,|_j = u^i\,|_j$$

and it is also helpful to introduce the following notation:

$$u_i\,|^j = u_i\,|_m\;g^{mj}, \quad u^i\,|^j = u^i\,|_m\;g^{mj}.$$

The divergence and curl can then be written as {▲Problem 10}

$$\begin{aligned}\mathrm{div}\mathbf{u} &= u^i\,|_i = u_i\,|^i\\ \mathrm{curl}\mathbf{u} &= e^{ijk}u_j\,|_i\;\mathbf{g}_k = e_{ijk}u^j\,|^i\;\mathbf{g}^k\end{aligned}.$$

Generalising Tensor Calculus from Cartesian to Curvilinear Coordinates

It was seen in §1.16.7 how formulae could be generalised from the Cartesian system to the corresponding formulae in curvilinear coordinates. In addition, formulae for the gradient, divergence and curl of tensor fields may be generalised to curvilinear components simply by replacing the partial derivatives with the covariant derivatives. Thus:

		Cartesian	Curvilinear
Gradient	Of a scalar field	$\mathrm{grad}\phi,\ \nabla\phi = \partial\phi / \partial x_i$	$\phi_{,i} = \phi\,\vert_i \equiv \partial\phi / \partial\Theta^i$
	of a vector field	$\mathrm{grad}\mathbf{u} = \partial u_i / \partial x_j$	$u^i\vert_j$
	of a tensor field	$\mathrm{grad}\mathbf{T} = \partial T_{ij} / \partial x_k$	$T^{ij}\vert_k$
Divergence	of a vector field	$\mathrm{div}\mathbf{u},\ \nabla\cdot\mathbf{u} = \partial u_i / \partial x_i$	$u^i\vert_i$
	of a tensor field	$\mathrm{div}\mathbf{T} = \partial T_{ij} / \partial x_j$	$T^{ij}\vert_j$
Curl	of a vector field	$\mathrm{curl}\mathbf{u},\ \nabla\times\mathbf{u} = \varepsilon_{ijk}\partial u_j / \partial x_i$	$e^{ijk} u_j\,\vert_i$

Table 1.18.1: generalising formulae from Cartesian to General Curvilinear Coordinates

All the tensor identities derived for Cartesian bases (§1.6.9, §1.14.3) hold also for curvilinear coordinates, for example {▲Problem 11}

$$
\begin{aligned}
\mathrm{grad}(\alpha\mathbf{v}) &= \alpha\,\mathrm{grad}\mathbf{v} + \mathbf{v}\otimes\mathrm{grad}\alpha \\
\mathrm{div}(\mathbf{vA}) &= \mathbf{v}\cdot\mathrm{div}\mathbf{A} + \mathbf{A} : \mathrm{grad}\mathbf{v}
\end{aligned}
$$

1.18.4 Partial Derivatives with respect to a Tensor

The notion of differentiation of one tensor with respect to another can be generalised from the Cartesian differentiation discussed in §1.15. For example:

$$
\begin{aligned}
\frac{\partial\Phi}{\partial\mathbf{A}} &= \frac{\partial\Phi}{\partial A_{ij}}\mathbf{g}_i\otimes\mathbf{g}_j = \frac{\partial\Phi}{\partial A_i^{\cdot j}}\mathbf{g}_i\otimes\mathbf{g}^j = \cdots \\
\frac{\partial\mathbf{B}}{\partial\mathbf{A}} &= \frac{\partial B_{ij}}{\partial A_{mn}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_m\otimes\mathbf{g}_n = \cdots \\
\frac{\partial\mathbf{A}}{\partial\mathbf{A}} &= \frac{\partial A_{ij}}{\partial A_{mn}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_m\otimes\mathbf{g}_n = \delta_i^m\delta_j^n\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_m\otimes\mathbf{g}_n \\
&= \mathbf{g}^m\otimes\mathbf{g}^n\otimes\mathbf{g}_m\otimes\mathbf{g}_n
\end{aligned}
$$

1.18.5 Orthogonal Curvilinear Coordinates

This section is based on the groundwork carried out in §1.16.9. In orthogonal curvilinear systems, it is best to write all equations in terms of the covariant base vectors, or in terms of the corresponding physical components, using the identities (see Eqn. 1.16.45)

$$\mathbf{g}^i = \frac{1}{h_i^2}\mathbf{g}_i = \frac{1}{h_i}\hat{\mathbf{g}}_i \qquad \text{(no sum)} \tag{1.18.30}$$

The Gradient of a Scalar Field

From the definition 1.18.23 for the gradient of a scalar field, and Eqn. 1.18.30, one has for an orthogonal curvilinear coordinate system,

$$\begin{aligned}\nabla\Phi &= \frac{1}{h_1^2}\frac{\partial\Phi}{\partial\Theta^1}\mathbf{g}_1 + \frac{1}{h_2^2}\frac{\partial\Phi}{\partial\Theta^2}\mathbf{g}_2 + \frac{1}{h_3^2}\frac{\partial\Phi}{\partial\Theta^3}\mathbf{g}_3 \\ &= \frac{1}{h_1}\frac{\partial\Phi}{\partial\Theta^1}\hat{\mathbf{g}}_1 + \frac{1}{h_2}\frac{\partial\Phi}{\partial\Theta^2}\hat{\mathbf{g}}_2 + \frac{1}{h_3}\frac{\partial\Phi}{\partial\Theta^3}\hat{\mathbf{g}}_3\end{aligned} \tag{1.18.31}$$

The Christoffel Symbols

The Christoffel symbols simplify considerably in orthogonal coordinate systems. First, from the definition 1.18.4,

$$\Gamma_{ij}^k = \frac{1}{h_k^2}\frac{\partial\mathbf{g}_i}{\partial\Theta^j}\cdot\mathbf{g}_k \tag{1.18.32}$$

Note that the introduction of the scale factors h into this and the following equations disrupts the summation and index notation convention used hitherto. To remain consistent, one should use the metric coefficients and leave this equation in the form

$$\Gamma_{ij}^k = \frac{\partial\mathbf{g}_i}{\partial\Theta^j}\cdot g^{km}\mathbf{g}_m$$

Now

$$\frac{\partial}{\partial\Theta^j}(\mathbf{g}_i\cdot\mathbf{g}_i) = 2\left(\frac{\partial\mathbf{g}_i}{\partial\Theta^j}\cdot\mathbf{g}_i\right) = 2h_i^2\Gamma_{ij}^i$$

and $\mathbf{g}_i\cdot\mathbf{g}_i = h_i^2$ so, in terms of the derivatives of the scale factors,

$$\Gamma_{ij}^i = \Gamma_{ij}^k\Big|_{k=i} = \frac{1}{h_i}\frac{\partial h_i}{\partial\Theta^j} \qquad \text{(no sum)} \tag{1.18.33}$$

Similarly, it can be shown that {▲Problem 14}

$$h_k^2\Gamma_{ij}^k = -h_i^2\Gamma_{jk}^i = -h_j^2\Gamma_{ki}^j = h_i^2\Gamma_{jk}^i \text{ when } i\neq j\neq k \tag{1.18.34}$$

so that the Christoffel symbols are zero when the indices are distinct, so that there are only 21 non-zero symbols of the 27. Further, {▲Problem 15}

$$\Gamma_{ii}^{k} = \Gamma_{ij}^{k}\Big|_{i=j} = -\frac{h_i}{h_k^2}\frac{\partial h_i}{\partial \Theta^k}, \quad i \neq k \quad \text{(no sum)} \tag{1.18.35}$$

From the symmetry condition (see Eqn. 1.18.4), only 15 of the 21 non-zero symbols are distinct:

$$\begin{gathered}
\Gamma_{11}^{1}, \Gamma_{12}^{1} = \Gamma_{21}^{1}, \Gamma_{13}^{1} = \Gamma_{31}^{1}, \Gamma_{22}^{1}, \Gamma_{33}^{1} \\
\Gamma_{11}^{2}, \Gamma_{12}^{2} = \Gamma_{21}^{2}, \Gamma_{22}^{2}, \Gamma_{23}^{2} = \Gamma_{32}^{2}, \Gamma_{33}^{2} \\
\Gamma_{11}^{3}, \Gamma_{13}^{3} = \Gamma_{31}^{3}, \Gamma_{22}^{3}, \Gamma_{23}^{3} = \Gamma_{32}^{3}, \Gamma_{33}^{3}
\end{gathered}$$

Note also that these are related to each other through the relation between (1.18.33, 1.18.35), i.e.

$$\Gamma_{ii}^{k} = -\frac{h_i^2}{h_k^2}\Gamma_{ik}^{i}, \quad i \neq k \quad \text{(no sum)}$$

so that

$$\begin{aligned}
&\Gamma_{11}^{1}, \quad \Gamma_{22}^{2}, \quad \Gamma_{33}^{3} \\
&\Gamma_{12}^{1} = \Gamma_{21}^{1} = -\frac{h_2^2}{h_1^2}\Gamma_{11}^{2}, \quad \Gamma_{13}^{1} = \Gamma_{31}^{1} = -\frac{h_3^2}{h_1^2}\Gamma_{11}^{3}, \quad \Gamma_{12}^{2} = \Gamma_{21}^{2} = -\frac{h_1^2}{h_2^2}\Gamma_{22}^{1} \\
&\Gamma_{23}^{2} = \Gamma_{32}^{2} = -\frac{h_3^2}{h_2^2}\Gamma_{22}^{3}, \quad \Gamma_{13}^{3} = \Gamma_{31}^{3} = -\frac{h_1^2}{h_3^2}\Gamma_{33}^{1}, \quad \Gamma_{23}^{3} = \Gamma_{32}^{3} = -\frac{h_2^2}{h_3^2}\Gamma_{33}^{2}
\end{aligned} \tag{1.18.36}$$

The Gradient of a Vector

From the definition 1.18.25, the gradient of a vector is

$$\operatorname{grad}\mathbf{v} = v^i\big|_j \mathbf{g}_i \otimes \mathbf{g}^j = \frac{1}{h_j^2} v^i\big|_j \mathbf{g}_i \otimes \mathbf{g}_j \quad \text{(no sum over } h_j\text{)} \tag{1.18.37}$$

In terms of physical components,

$$\begin{aligned}
\operatorname{grad}\mathbf{v} &= \frac{1}{h_j^2}\left(\frac{\partial v^i}{\partial \Theta^j} + v^k \Gamma_{kj}^{i}\right)\mathbf{g}_i \otimes \mathbf{g}_j \\
&= \frac{1}{h_j}\left(\frac{\partial v^{\langle i\rangle}}{\partial \Theta^j} - \Gamma_{ij}^{i} v^{\langle i\rangle} + \frac{h_i}{h_k}\Gamma_{kj}^{i} v^{\langle k\rangle}\right)\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}_j
\end{aligned} \tag{1.18.38}$$

The Divergence of a Vector

From the definition 1.18.27, the divergence of a vector is $\operatorname{div}\mathbf{v} = v^i\big|_i$ or {▲Problem 16}

$$\operatorname{div}\mathbf{v} = \frac{\partial v_i}{\partial \Theta^i} + v_k \Gamma^i_{ki} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial\left(v^{\langle 1\rangle} h_2 h_3\right)}{\partial \Theta^1} + \frac{\partial\left(v^{\langle 2\rangle} h_1 h_3\right)}{\partial \Theta^2} + \frac{\partial\left(v^{\langle 3\rangle} h_1 h_2\right)}{\partial \Theta^3}\right] \tag{1.18.39}$$

The Curl of a Vector

From §1.16.10, the permutation symbol in orthogonal curvilinear coordinates reduceS to

$$e^{ijk} = \frac{1}{h_1 h_2 h_3}\varepsilon^{ijk} \tag{1.18.40}$$

where $\varepsilon^{ijk} = \varepsilon_{ijk}$ is the Cartesian permutation symbol. From the definition 1.18.29, the curl of a vector is then

$$\begin{aligned}\operatorname{curl}\mathbf{v} &= \frac{1}{h_1 h_2 h_3}\varepsilon_{ijk}\frac{\partial v_j}{\partial \Theta^i}\mathbf{g}_k = \frac{1}{h_1 h_2 h_3}\left\{\left[\frac{\partial v_2}{\partial \Theta^1} - \frac{\partial v_1}{\partial \Theta^2}\right]\mathbf{g}_3 + \cdots\right\} \\ &= \frac{1}{h_1 h_2 h_3}\left\{\left[\frac{\partial\left(v^2 h_2^2\right)}{\partial \Theta^1} - \frac{\partial\left(v^1 h_1^2\right)}{\partial \Theta^2}\right]\mathbf{g}_3 + \cdots\right\} \\ &= \frac{1}{h_1 h_2 h_3}\left\{\left[\frac{\partial\left(v^{\langle 2\rangle} h_2\right)}{\partial \Theta^1} - \frac{\partial\left(v^{\langle 1\rangle} h_1\right)}{\partial \Theta^2}\right]h_3\hat{\mathbf{g}}_3 + \cdots\right\}\end{aligned} \tag{1.18.41}$$

or

$$\operatorname{curl}\mathbf{v} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} h_1\hat{\mathbf{g}}_1 & h_2\hat{\mathbf{g}}_2 & h_3\hat{\mathbf{g}}_3 \\ \dfrac{\partial}{\partial \Theta^1} & \dfrac{\partial}{\partial \Theta^2} & \dfrac{\partial}{\partial \Theta^3} \\ h_1 v^{\langle 1\rangle} & h_2 v^{\langle 2\rangle} & h_3 v^{\langle 3\rangle}\end{vmatrix} \tag{1.18.42}$$

The Laplacian

From the above results, the Laplacian is given by

$$\nabla^2\Phi = \nabla\cdot\nabla\Phi = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial \Theta^1}\left(\frac{h_2 h_3}{h_1}\frac{\partial \Phi}{\partial \Theta^1}\right) + \frac{\partial}{\partial \Theta^2}\left(\frac{h_1 h_3}{h_2}\frac{\partial \Phi}{\partial \Theta^2}\right) + \frac{\partial}{\partial \Theta^3}\left(\frac{h_1 h_2}{h_3}\frac{\partial \Phi}{\partial \Theta^3}\right)\right]$$

Divergence of a Tensor

From the definition 1.18.28, and using 1.16.59, 1.16.62 {▲Problem 17}

$$\begin{aligned}\operatorname{div}\mathbf{A} = A_i^{\cdot j}\,|_j\ \mathbf{g}^i &= \left\{\frac{\partial A_i^{\cdot j}}{\partial \Theta^j} + \Gamma_{mj}^j A_i^{\cdot m} - \Gamma_{ij}^m A_m^{\cdot j}\right\}\mathbf{g}^i \\ &= \left\{\frac{1}{h_i}\frac{\partial}{\partial \Theta^j}\left(\frac{h_i}{h_j}A^{\langle ij\rangle}\right) + \frac{1}{h_m}\Gamma_{mj}^j A^{\langle im\rangle} - \frac{h_m}{h_i h_j}\Gamma_{ij}^m A^{\langle mj\rangle}\right\}\hat{\mathbf{g}}_i\end{aligned} \tag{1.18.43}$$

Examples

1. Cylindrical Coordinates

Gradient of a Scalar Field:

$$\nabla\Phi = \frac{\partial \Phi}{\partial \Theta^1}\hat{\mathbf{g}}_1 + \frac{1}{\Theta^2}\frac{\partial \Phi}{\partial \Theta_2}\hat{\mathbf{g}}_2 + \frac{\partial \Phi}{\partial \Theta^3}\hat{\mathbf{g}}_3$$

Christoffel symbols:

With $h_1 = 1, h_2 = \Theta^1, h_3 = 1$, there are two distinct non-zero symbols:

$$\Gamma_{22}^1 = -\Theta^1$$

$$\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{\Theta^1}$$

Derivatives of the base vectors:

The non-zero derivatives are

$$\frac{\partial \mathbf{g}_1}{\partial \Theta^2} = \frac{\partial \mathbf{g}_2}{\partial \Theta^1} = \frac{1}{\Theta^1}\mathbf{g}_2, \quad \frac{\partial \mathbf{g}_2}{\partial \Theta^2} = -\Theta^1\mathbf{g}_1$$

and in terms of physical components, the non-zero derivatives are

$$\frac{\partial \hat{\mathbf{g}}_1}{\partial \Theta^2} = \hat{\mathbf{g}}_2, \quad \frac{\partial \hat{\mathbf{g}}_2}{\partial \Theta^2} = -\hat{\mathbf{g}}_1$$

which agree with 1.6.32.

The Divergence (see 1.6.33), Curl (see 1.6.34) and Gradient {▲Problem 18} (see 1.14.18) of a vector:

$$\operatorname{div}\mathbf{v} = \left[\frac{\partial v_{\langle 1\rangle}}{\partial \Theta^1} + \frac{v_{\langle 1\rangle}}{\Theta^1} + \frac{1}{\Theta^1}\frac{\partial v_{\langle 2\rangle}}{\partial \Theta^2} + \frac{\partial v_{\langle 3\rangle}}{\partial \Theta^3}\right]$$

$$\operatorname{curl}\mathbf{v} = \frac{1}{\Theta^1}\begin{vmatrix}\hat{\mathbf{g}}_1 & \Theta^1\hat{\mathbf{g}}_2 & \hat{\mathbf{g}}_3 \\ \dfrac{\partial}{\partial \Theta^1} & \dfrac{\partial}{\partial \Theta^2} & \dfrac{\partial}{\partial \Theta^3} \\ v^{\langle 1\rangle} & \Theta^1 v^{\langle 2\rangle} & v^{\langle 3\rangle}\end{vmatrix}$$

$$\operatorname{grad}\mathbf{v} = \frac{\partial v_{\langle 1\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_1 + \frac{\partial v_{\langle 2\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_1 + \frac{\partial v_{\langle 3\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_1$$
$$+\frac{1}{\Theta^1}\left(\frac{\partial v_{\langle 1\rangle}}{\partial\Theta^2} - v_{\langle 2\rangle}\right)\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_2 + \frac{1}{\Theta^1}\left(\frac{\partial v_{\langle 2\rangle}}{\partial\Theta^2} + v_{\langle 1\rangle}\right)\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_2$$
$$+\frac{1}{\Theta^1}\left(\frac{\partial v_{\langle 3\rangle}}{\partial\Theta^2}\right)\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_2 + \frac{\partial v_{\langle 1\rangle}}{\partial\Theta^3}\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_3 + \frac{\partial v_{\langle 2\rangle}}{\partial\Theta^3}\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_3 + \frac{\partial v_{\langle 3\rangle}}{\partial\Theta^3}\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_3$$

The Divergence of a tensor {▲Problem 19} (see 1.14.19):

$$\operatorname{div}\mathbf{A} = \left\{\frac{\partial A^{\langle 11\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 12\rangle}}{\partial\Theta^2} + \frac{\partial A^{\langle 13\rangle}}{\partial\Theta^3} + \frac{A^{\langle 11\rangle} - A^{\langle 22\rangle}}{\Theta^1}\right\}\hat{\mathbf{g}}_1$$
$$+\left\{\frac{\partial A^{\langle 21\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 22\rangle}}{\partial\Theta^2} + \frac{\partial A^{\langle 23\rangle}}{\partial\Theta^3} + \frac{A^{\langle 21\rangle} + A^{\langle 12\rangle}}{\Theta^1}\right\}\hat{\mathbf{g}}_2$$
$$+\left\{\frac{\partial A^{\langle 31\rangle}}{\partial\Theta^1} + \frac{A^{\langle 31\rangle}}{\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 32\rangle}}{\partial\Theta^2} + \frac{\partial A^{\langle 33\rangle}}{\partial\Theta^3}\right\}\hat{\mathbf{g}}_3$$

2. Spherical Coordinates

Gradient of a Scalar Field:

$$\nabla\Phi = \frac{\partial\Phi}{\partial\Theta^1}\hat{\mathbf{g}}_1 + \frac{1}{\Theta^1}\frac{\partial\Phi}{\partial\Theta^2}\hat{\mathbf{g}}_2 + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial\Phi}{\partial\Theta^3}\hat{\mathbf{g}}_3$$

Christoffel symbols:

With $h_1 = 1,\ h_2 = \Theta^1,\ h_3 = \Theta^1\sin\Theta^2$, there are six distinct non-zero symbols:

$$\Gamma^1_{22} = -\Theta^1, \Gamma^1_{33} = -\Theta^1\sin^2\Theta^2$$
$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{\Theta^1}, \Gamma^2_{33} = -\sin\Theta^2\cos\Theta^2$$
$$\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{\Theta^1},\ \Gamma^3_{23} = \Gamma^3_{32} = \cot\Theta^2$$

Derivatives of the base vectors:

The non-zero derivatives are

$$\frac{\partial\mathbf{g}_1}{\partial\Theta^2} = \frac{\partial\mathbf{g}_2}{\partial\Theta^1} = \frac{1}{\Theta^1}\mathbf{g}_2,\quad \frac{\partial\mathbf{g}_1}{\partial\Theta^3} = \frac{\partial\mathbf{g}_3}{\partial\Theta^1} = \frac{1}{\Theta^1}\mathbf{g}_3,\quad \frac{\partial\mathbf{g}_2}{\partial\Theta^2} = -\Theta^1\mathbf{g}_1$$
$$\frac{\partial\mathbf{g}_2}{\partial\Theta_3} = \frac{\partial\mathbf{g}_3}{\partial\Theta_2} = \cot\Theta^2\mathbf{g}_3,\quad \frac{\partial\mathbf{g}_3}{\partial\Theta_3} = -\Theta^1\sin^2\Theta^2\mathbf{g}_1 - \sin\Theta^2\cos\Theta^2\mathbf{g}_2$$

and in terms of physical components, the non-zero derivatives are

$$\frac{\partial\hat{\mathbf{g}}_1}{\partial\Theta^2} = \hat{\mathbf{g}}_2,\quad \frac{\partial\hat{\mathbf{g}}_1}{\partial\Theta^3} = \sin\Theta^2\hat{\mathbf{g}}_3,\quad \frac{\partial\hat{\mathbf{g}}_2}{\partial\Theta^2} = -\hat{\mathbf{g}}_1$$
$$\frac{\partial\hat{\mathbf{g}}_2}{\partial\Theta^3} = \cos\Theta^2\hat{\mathbf{g}}_3,\quad \frac{\partial\hat{\mathbf{g}}_3}{\partial\Theta^3} = -\sin\Theta^2\hat{\mathbf{g}}_1 - \cos\Theta^2\hat{\mathbf{g}}_2$$

which agree with 1.6.37.

The Divergence (see 1.6.38), Curl and Gradient of a Vector:

$$\operatorname{div}\mathbf{v} = \frac{1}{\left(\Theta^1\right)^2}\frac{\partial\left(\left(\Theta^1\right)^2 v_{\langle 1\rangle}\right)}{\partial\Theta^1} + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial\left(\sin\Theta^2 v_{\langle 2\rangle}\right)}{\partial\Theta^2} + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial v_{\langle 3\rangle}}{\partial\Theta^3}$$

$$\operatorname{curl}\mathbf{v} = \frac{1}{\left(\Theta^1\right)^2\sin\Theta^2}\begin{vmatrix} \hat{\mathbf{g}}_1 & \Theta^1\hat{\mathbf{g}}_2 & \Theta^1\sin\Theta^2\hat{\mathbf{g}}_3 \\ \dfrac{\partial}{\partial\Theta^1} & \dfrac{\partial}{\partial\Theta^2} & \dfrac{\partial}{\partial\Theta^3} \\ v^{\langle 1\rangle} & \Theta^1 v^{\langle 2\rangle} & \Theta^1\sin\Theta^2 v^{\langle 3\rangle}\end{vmatrix}$$

$$\begin{aligned}\operatorname{grad}\mathbf{v} = &\frac{\partial v_{\langle 1\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_1 + \left(\frac{1}{\Theta^1}\frac{\partial v_{\langle 1\rangle}}{\partial\Theta^2} - \frac{v_{\langle 2\rangle}}{\Theta^1}\right)\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_2 \\ &+ \left(\frac{1}{\Theta^1\sin\Theta^2}\frac{\partial v_{\langle 1\rangle}}{\partial\Theta^3} - \frac{v_{\langle 3\rangle}}{\Theta^1}\right)\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_3 + \frac{\partial v_{\langle 2\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_1 \\ &+ \left(\frac{1}{\Theta^1}\frac{\partial v_{\langle 2\rangle}}{\partial\Theta^2} + \frac{v_{\langle 1\rangle}}{\Theta^1}\right)\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_2 + \left(\frac{1}{\Theta^1\sin\Theta^2}\frac{\partial v_{\langle 2\rangle}}{\partial\Theta^3} - \cot\Theta^2\frac{v_{\langle 3\rangle}}{\Theta^1}\right)\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_3 \\ &+ \frac{\partial v_{\langle 3\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_1 + \frac{1}{\Theta^1}\frac{\partial v_{\langle 3\rangle}}{\partial\Theta^2}\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_2 \\ &+ \left(\frac{1}{\Theta^1\sin\Theta^2}\frac{\partial v_{\langle 3\rangle}}{\partial\Theta^3} + \frac{v_{\langle 1\rangle}}{\Theta^1} + \cot\Theta^2\frac{v_{\langle 2\rangle}}{\Theta^1}\right)\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_3\end{aligned}$$

The Divergence of a tensor {▲Problem 20}

$$\begin{aligned}\operatorname{div}\mathbf{A} = &\left\{\frac{\partial A^{\langle 11\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 12\rangle}}{\partial\Theta^2} + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial A^{\langle 13\rangle}}{\partial\Theta^3} + \frac{2A^{\langle 11\rangle} + \cot\Theta^2 A^{\langle 12\rangle} - A^{\langle 22\rangle} - A^{\langle 33\rangle}}{\Theta^1}\right\}\hat{\mathbf{g}}_1 \\ &+ \left\{\frac{\partial A^{\langle 21\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 22\rangle}}{\partial\Theta^2} + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial A^{\langle 23\rangle}}{\partial\Theta^3} + \frac{A^{\langle 12\rangle} + 2A^{\langle 21\rangle} + \cot\Theta^2\left(A^{\langle 22\rangle} - A^{\langle 33\rangle}\right)}{\Theta^1}\right\}\hat{\mathbf{g}}_2 \\ &+ \left\{\frac{\partial A^{\langle 31\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial A^{\langle 32\rangle}}{\partial\Theta^2} + \frac{1}{\Theta^1\sin\Theta^2}\frac{\partial A^{\langle 33\rangle}}{\partial\Theta^3} + \frac{A^{\langle 13\rangle} + 2A^{\langle 31\rangle} + \cot\Theta^2\left(A^{\langle 23\rangle} + A^{\langle 32\rangle}\right)}{\Theta^1}\right\}\hat{\mathbf{g}}_3\end{aligned}$$

1.18.6 Problems

1 Show that the Christoffel symbol of the second kind is symmetric, i.e. $\Gamma_{ij}^k = \Gamma_{ji}^k$, and that it is explicitly given by $\Gamma_{ij}^k = \dfrac{\partial\mathbf{g}_i}{\partial\Theta^j}\cdot\mathbf{g}^k$.

2 Consider the scalar-valued function $\Phi = (\mathbf{A}\mathbf{u})\cdot\mathbf{v} = A_{ij}u^i v^j$. By taking the gradient of this function, and using the relation for the covariant derivative of **A**, i.e. $A_{ij}|_k = A_{ij,k} - \Gamma_{ik}^m A_{mj} - \Gamma_{jk}^m A_{im}$, show that

$$\frac{\partial\left(A_{ij}u^i v^j\right)}{\partial\Theta^k} = \left(A_{ij}u^i v^j\right)|_k,$$

i.e. the partial derivative and covariant derivative are equivalent for a scalar-valued function.

3 Prove 1.18.9:

(i) $\dfrac{\partial g_{ij}}{\partial \Theta^k} = \Gamma_{ikj} + \Gamma_{jki}$, (ii) $\dfrac{\partial g^{ij}}{\partial \Theta^k} = -g^{im}\Gamma^j_{km} - g^{jm}\Gamma^i_{km}$

[Hint: for (ii), first differentiate Eqn. 1.16.10, $g^{ij} g_{kj} = \delta^i_k$.]

4 Derive 1.18.13, relating the Christoffel symbols to the partial derivatives of $\sqrt{g}$ and $\log\left(\sqrt{g}\right)$. [Hint: begin by using the chain rule $\dfrac{\partial g}{\partial \Theta^j} = \dfrac{\partial g}{\partial g_{mn}} \dfrac{\partial g_{mn}}{\partial \Theta^j}$.]

5 Use the definition of the covariant derivative of second order tensor components, Eqn. 1.18.18, to show that (i) $g_{ij}|_k = 0$ and (ii) $g^{ij}|_k = 0$.

6 Use the definition of the gradient of a vector, 1.18.25, to show that $\operatorname{div}\mathbf{u} = \operatorname{grad}\mathbf{u} : \mathbf{I} = u^i \,|_i$.

7 Derive the expression $\operatorname{div}\mathbf{u} = \left(1/\sqrt{g}\right)\partial\left(\sqrt{g}u^i\right)/\partial\Theta^i$

8 Use 1.16.54 to show that $\mathbf{g}^k \times \left(\partial \mathbf{u}/\partial\Theta^k\right) = e^{ijk} u_j \,|_i \,\mathbf{g}_k$.

9 Use the relation $\varepsilon^{ijk}\varepsilon_{imn} = \delta^j_m\delta^k_n - \delta^k_m\delta^j_n$ (see Eqn. 1.3.19) to show that

$$\operatorname{curl}(\mathbf{u}\times\mathbf{v}) = \begin{vmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 \\ \dfrac{\partial}{\partial\Theta^1} + \Gamma^k_{1k} & \dfrac{\partial}{\partial\Theta^2} + \Gamma^k_{2k} & \dfrac{\partial}{\partial\Theta^3} + \Gamma^k_{3k} \\ \left(u^2v^3 - u^3v^2\right) & -\left(u^1v^3 - u^3v^1\right) & \left(u^1v^2 - u^2v^1\right) \end{vmatrix}.$$

10 Show that (i) $u^i \,|_i = u_i \,|^i$, (ii) $e^{ijk} u_j \,|_i \,\mathbf{g}_k = e_{ijk} u^j \,|^i \,\mathbf{g}^k$

11 Show that

(i) $\operatorname{grad}(\alpha\mathbf{v}) = \alpha\operatorname{grad}\mathbf{v} + \mathbf{v}\otimes\operatorname{grad}\alpha$, (ii) $\operatorname{div}(\mathbf{vA}) = \mathbf{v}\cdot\operatorname{div}\mathbf{A} + \mathbf{A} : \operatorname{grad}\mathbf{v}$

[Hint: you might want to use the relation $\mathbf{aT}\cdot\mathbf{b} = \mathbf{T} : (\mathbf{a}\otimes\mathbf{b})$ for the second of these.]

12 Derive the relation $\partial(\operatorname{tr}\mathbf{A})/\partial\mathbf{A} = \mathbf{I}$ in curvilinear coordinates.

13 Consider a (two dimensional) curvilinear coordinate system with covariant base vectors $\mathbf{g}_1 = \Theta^2\mathbf{e}_1 - 2\mathbf{e}_2$, $\mathbf{g}_2 = \Theta^1\mathbf{e}_1$.

(a) Evaluate the transformation equations $x^i = x^i(\Theta^j)$ and the Jacobian *J*.

(b) Evaluate the inverse transformation equations $\Theta^i = \Theta^i\left(x^j\right)$ and the contravariant base vectors $\mathbf{g}^i$.

(c) Evaluate the metric coefficients g_{ij}, g^{ij} and the function *g*:

(d) Evaluate the Christoffel symbols (only 2 are non-zero

(e) Consider the scalar field $\Phi = \Theta^1 + \Theta^2$. Evaluate $\operatorname{grad}\Phi$.

(f) Consider the vector fields $\mathbf{u} = \mathbf{g}_1 + \Theta^2\mathbf{g}_2$, $\mathbf{v} = -\left(\Theta^1\right)^2\mathbf{g}_1 + 2\mathbf{g}_2$:

(i) Evaluate the covariant components of the vectors **u** and **v**

(ii) Evaluate $\operatorname{div}\mathbf{u}$, $\operatorname{div}\mathbf{v}$

(iii)Evaluate $\operatorname{curl}\mathbf{u}$, $\operatorname{curl}\mathbf{v}$

(iv)Evaluate $\operatorname{grad}\mathbf{u}$, $\operatorname{grad}\mathbf{v}$

(g) Verify the vector identities

$$\operatorname{div}(\Phi\mathbf{u}) = \Phi \operatorname{div}\mathbf{u} + \operatorname{grad}\Phi \cdot \mathbf{u}$$
$$\operatorname{curl}(\Phi\mathbf{u}) = \Phi \operatorname{curl}\mathbf{u} + \operatorname{grad}\Phi \times \mathbf{u}$$
$$\operatorname{div}(\mathbf{u}\times\mathbf{v}) = \mathbf{v}\cdot\operatorname{curl}\mathbf{u} - \mathbf{u}\cdot\operatorname{curl}\mathbf{v}$$
$$\operatorname{curl}(\operatorname{grad}\Phi) = \mathbf{o}$$
$$\operatorname{div}(\operatorname{curl}\mathbf{u}) = 0$$

(h) Verify the identities

$$\operatorname{grad}(\Phi\mathbf{v}) = \Phi\operatorname{grad}\mathbf{v} + \mathbf{v}\otimes\operatorname{grad}\Phi$$
$$\operatorname{grad}(\mathbf{u}\cdot\mathbf{v}) = (\operatorname{grad}\mathbf{u})^{\mathrm{T}}\mathbf{v} + (\operatorname{grad}\mathbf{v})^{\mathrm{T}}\mathbf{u}$$
$$\operatorname{div}(\mathbf{u}\otimes\mathbf{v}) = (\operatorname{grad}\mathbf{u})\mathbf{v} + (\operatorname{div}\mathbf{v})\mathbf{u}$$
$$\operatorname{curl}(\mathbf{u}\times\mathbf{v}) = \mathbf{u}\operatorname{div}\mathbf{v} - \mathbf{v}\operatorname{div}\mathbf{u} + (\operatorname{grad}\mathbf{u})\mathbf{v} - (\operatorname{grad}\mathbf{v})\mathbf{u}$$

(i) Consider the tensor field

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & -\Theta^2 \end{bmatrix} \left(\mathbf{g}^i \otimes \mathbf{g}^j\right)$$

Evaluate all contravariant and mixed components of the tensor **A**

14 Use the fact that $\mathbf{g}_k \cdot \mathbf{g}_i = 0$, $k \neq i$ to show that $h_k^2 \Gamma_{ij}^k = -h_i^2 \Gamma_{jk}^{i}{}^{k}_{ij}$. Then permutate the indices to show that $h_k^2 \Gamma_{ij}^k = -h_i^2 \Gamma_{jk}^i = -h_j^2 \Gamma_{ki}^j = h_i^2 \Gamma_{jk}^i$ when $i \neq j \neq k$.

15 Use the relation

$$\frac{\partial}{\partial\Theta^i}\left(\mathbf{g}_i \cdot \mathbf{g}_j\right) = 0, \quad i \neq j$$

to derive $\Gamma_{ii}^j = -\dfrac{h_i^2}{h_j^2}\Gamma_{ij}^i$.

16 Derive the expression 1.18.39 for the divergence of a vector field **v**.

17 Derive 1.18.43 for the divergence of a tensor in orthogonal coordinate systems.

18 Use the expression 1.18.38 to derive the expression for the gradient of a vector field in cylindrical coordinates.

19 Use the expression 1.18.43 to derive the expression for the divergence of a tensor field in cylindrical coordinates.

20 Use the expression 1.18.43 to derive the expression for the divergence of a tensor field in spherical coordinates.

1.19 Curvilinear Coordinates: Curved Geometries

In this section is examined the special case of a two-dimensional curved surface.

1.19.1 Monoclinic Coordinate Systems

Base Vectors

A curved surface can be defined using two covariant base vectors $\mathbf{a}_1, \mathbf{a}_2$, with the third base vector, $\mathbf{a}_3$, everywhere of unit size and normal to the other two, Fig. 1.19.1 These base vectors form a **monoclinic** reference frame, that is, only one of the angles between the base vectors is not necessarily a right angle.

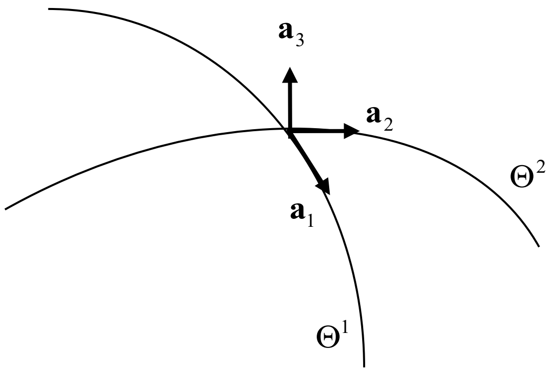


Figure 1.19.1: Geometry of the Curved Surface

In what follows, in the index notation, Greek letters such as α, β take values 1 and 2; as before, Latin letters take values from 1..3.

Since $\mathbf{a}^3 = \mathbf{a}_3$ and

$$a_{\alpha 3} = \mathbf{a}_\alpha \cdot \mathbf{a}_3 = 0, \qquad a^{\alpha 3} = \mathbf{a}^\alpha \cdot \mathbf{a}^3 = 0 \tag{1.19.1}$$

the determinant of metric coefficients is

$$J^2 = \begin{vmatrix} g_{11} & g_{12} & 0 \\ g_{21} & g_{22} & 0 \\ 0 & 0 & 1 \end{vmatrix}, \qquad \frac{1}{J^2} = \begin{vmatrix} g^{11} & g^{12} & 0 \\ g^{21} & g^{22} & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{1.19.2}$$

The Cross Product

Particularising the results of §1.16.10, define the surface permutation symbol to be the triple scalar product

$$e_{\alpha\beta} \equiv \mathbf{a}_\alpha \cdot \mathbf{a}_\beta \times \mathbf{a}_3 = \varepsilon_{\alpha\beta}\sqrt{g}, \qquad e^{\alpha\beta} \equiv \mathbf{a}^\alpha \cdot \mathbf{a}^\beta \times \mathbf{a}^3 = \varepsilon^{\alpha\beta}\frac{1}{\sqrt{g}} \tag{1.19.3}$$

where $\varepsilon_{\alpha\beta} = \varepsilon^{\alpha\beta}$ is the Cartesian permutation symbol, $\varepsilon_{12} = +1$, $\varepsilon_{21} = -1$, and zero otherwise, with

$$e^{\alpha\beta}e_{\mu\eta} = \varepsilon^{\alpha\beta}\varepsilon_{\mu\eta}, \qquad e^{\alpha\beta}e_{\mu\eta} = \delta^\alpha_\mu\delta^\beta_\eta - \delta^\beta_\mu\delta^\alpha_\eta = e^{\beta\alpha}e_{\eta\mu} \tag{1.19.4}$$

From 1.19.3,

$$\begin{aligned} \mathbf{a}_\alpha \times \mathbf{a}_\beta &= e_{\alpha\beta}\mathbf{a}^3 \\ \mathbf{a}^\alpha \times \mathbf{a}^\beta &= e^{\alpha\beta}\mathbf{a}_3 \end{aligned} \tag{1.19.5}$$

and so

$$\mathbf{a}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\sqrt{g}} \tag{1.19.6}$$

The cross product of surface vectors, that is, vectors with component in the normal ($\mathbf{g}_3$) direction zero, can be written as

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= e_{\alpha\beta}u^\alpha v^\beta \mathbf{a}^3 = \sqrt{g}\begin{vmatrix} u^1 & u^2 \\ v^1 & v^2 \end{vmatrix}\mathbf{a}^3 \\ &= e^{\alpha\beta}u_\alpha v_\beta \mathbf{a}_3 = \frac{1}{\sqrt{g}}\begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}\mathbf{a}_3 \end{aligned} \tag{1.19.7}$$

The Metric and Surface elements

Considering a line element lying within the surface, so that $\Theta^3 = 0$, the metric for the surface is

$$(\Delta s)^2 = d\mathbf{s} \cdot d\mathbf{s} = \left(d\Theta^\alpha \mathbf{a}_\alpha\right) \cdot \left(d\Theta^\beta \mathbf{a}_\beta\right) = g_{\alpha\beta}d\Theta^\alpha d\Theta^\beta \tag{1.19.8}$$

which is in this context known as the **first fundamental form of the surface**.

Similarly, from 1.16.41, a surface element is given by

$$\Delta S = \sqrt{g}\Delta\Theta^1\Delta\Theta^2 \tag{1.19.9}$$

Christoffel Symbols

The Christoffel symbols can be simplified as follows. A differentiation of $\mathbf{a}_3 \cdot \mathbf{a}_3 = 1$ leads to

$$\mathbf{a}_{3,\alpha} \cdot \mathbf{a}_3 = -\mathbf{a}_{3,\alpha} \cdot \mathbf{a}_3 \tag{1.19.10}$$

so that, from Eqn 1.18.6,

$$\Gamma_{3\alpha 3} = \Gamma_{\alpha 33} = 0 \tag{1.19.11}$$

Further, since $\partial \mathbf{a}_3 / \partial \Theta^3 = 0$,

$$\Gamma_{33\alpha} = 0, \quad \Gamma_{333} = 0 \tag{1.19.12}$$

These last two equations imply that the Γ_{ijk} vanish whenever two or more of the subscripts are 3.

Next, differentiate 1.19.1 to get

$$\mathbf{a}_{\alpha,\beta} \cdot \mathbf{a}_3 = -\mathbf{a}_{3,\beta} \cdot \mathbf{a}_\alpha, \qquad \mathbf{a}^{\alpha}{}_{,\beta} \cdot \mathbf{a}^3 = -\mathbf{a}^3{}_{,\beta} \cdot \mathbf{a}^\alpha \tag{1.19.13}$$

and Eqns. 1.18.6 now lead to

$$\Gamma_{\alpha\beta 3} = \Gamma_{\beta\alpha 3} = -\Gamma_{3\beta\alpha} = -\Gamma_{\beta 3\alpha} \tag{1.19.14}$$

From 1.18.8, using 1.19.11,

$$\begin{aligned} \Gamma^3_{\alpha\beta} &= \Gamma_{\alpha\beta\gamma} g^{\gamma 3} + \Gamma_{\alpha\beta 3} g^{33} = \Gamma_{\alpha\beta 3} \\ \Gamma^3_{3\alpha} &= \Gamma_{3\alpha\beta} g^{\beta 3} + \Gamma_{3\alpha 3} g^{33} = \Gamma_{3\alpha 3} = 0 \end{aligned} \tag{1.19.15}$$

and, similarly {▲Problem 1}

$$\Gamma^3_{\alpha 3} = \Gamma^\alpha_{33} = \Gamma^3_{33} = 0 \tag{1.19.16}$$

1.19.2 The Curvature Tensor

In this section is introduced a tensor which, with the metric coefficients, completely describes the surface.

First, although the base vector $\mathbf{a}_3$ maintains unit length, its direction changes as a function of the coordinates Θ^1, Θ^2, and its derivative is, from 1.18.2 or 1.18.5 (and using 1.19.15)

$$\frac{\partial \mathbf{a}_3}{\partial \Theta^\alpha} = \Gamma^k_{3\alpha} \mathbf{a}_k = \Gamma^\beta_{3\alpha} \mathbf{a}_\beta, \qquad \frac{\partial \mathbf{a}^3}{\partial \Theta^\alpha} = -\Gamma^3_{\alpha k} \mathbf{a}^k = -\Gamma^3_{\alpha\beta} \mathbf{a}^\beta \tag{1.19.17}$$

Define now the **curvature tensor K** to have the covariant components $K_{\alpha\beta}$, through

$$\frac{\partial \mathbf{a}_3}{\partial \Theta^\alpha} = -K_{\alpha\beta}\mathbf{a}^\beta \tag{1.19.18}$$

and it follows from 1.19.13, 1.19.15a and 1.19.14,

$$K_{\alpha\beta} = \Gamma^3_{\alpha\beta} = \Gamma_{\alpha\beta 3} = -\Gamma_{3\beta\alpha} \tag{1.19.19}$$

and, since these Christoffel symbols are symmetric in the α, β, *the curvature tensor is symmetric*.

The mixed and contravariant components of the curvature tensor follows from 1.16.58-9:

$$\begin{gathered} K_\alpha^\beta = g^{\gamma\beta}K_{\alpha\gamma} = g_{\alpha\gamma}K^{\gamma\beta}, \quad K^{\alpha\beta} = g^{\alpha\gamma}g^{\beta\lambda}K_{\gamma\lambda} \\ \frac{\partial \mathbf{a}_3}{\partial \Theta^\alpha} \equiv -K_{\alpha\beta}\mathbf{a}^\beta = -K_{\alpha\beta}g^{\gamma\beta}\mathbf{a}_\gamma = -K_\alpha^\gamma \mathbf{a}_\gamma \end{gathered} \tag{1.19.20}$$

and the "dot" is not necessary in the mixed notation because of the symmetry property. From these and 1.18.8, it follows that

$$K_\beta^\alpha = g^{\gamma\alpha}K_{\gamma\beta} = -g^{\gamma\alpha}\Gamma_{3\beta\gamma} = -\Gamma^\alpha_{3\beta} = -\Gamma^\alpha_{\beta 3} \tag{1.19.21}$$

Also,

$$\begin{aligned} d\mathbf{a}_3 \cdot d\mathbf{s} &= \left(\mathbf{a}_{3,\alpha} d\Theta^\alpha\right)\cdot\left(d\Theta^\beta \mathbf{a}_\beta\right) \\ &= \left(-K_{\alpha\gamma} d\Theta^\alpha \mathbf{a}^\gamma\right)\cdot\left(d\Theta^\beta \mathbf{a}_\beta\right) \\ &= -K_{\alpha\beta} d\Theta^\alpha d\Theta^\beta \end{aligned} \tag{1.19.22}$$

which is known as the **second fundamental form of the surface**.

From 1.19.19 and the definitions of the Christoffel symbols, 1.18.4, 1.18.6, the curvature can be expressed as

$$K_{\alpha\beta} = \frac{\partial \mathbf{a}_\alpha}{\partial \Theta^\beta}\cdot \mathbf{a}_3 = -\frac{\partial \mathbf{a}_3}{\partial \Theta^\beta}\cdot \mathbf{a}_\alpha \tag{1.19.23}$$

showing that the curvature is a measure of the change of the base vector $\mathbf{a}_\alpha$ along the Θ^β curve, in the direction of the normal vector; alternatively, the rate of change of the normal vector along Θ^β, in the direction $-\mathbf{a}_\alpha$. Looking at this in more detail, consider now the change in the normal vector $\mathbf{a}_3$ in the Θ^1 direction, Fig. 1.19.2. Then

$$d\mathbf{a}_3 = \mathbf{a}_{3,1} d\Theta^1 = -K_1^\gamma d\Theta^1 \mathbf{a}_\gamma \tag{1.19.24}$$

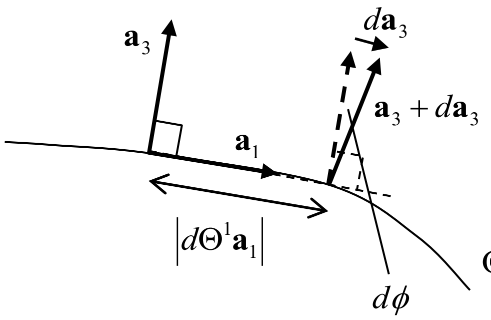


Figure 1.19.2: Curvature of the Surface

Taking the case of $K_1^1 \neq 0, K_1^2 = 0$, one has $d\mathbf{a}_3 = -K_1^1 d\Theta^1\mathbf{a}_1$. From Fig. 1.19.2, and since the normal vector is of unit length, the magnitude $\left|d\mathbf{a}_3\right|$ equals $d\phi$, the small angle through which the normal vector rotates as one travels along the Θ^1 coordinate curve. The **curvature** of the surface is defined to be the rate of change of the angle ϕ:[1]

$$\frac{d\phi}{ds} = \frac{\left|-K_1^1 d\Theta^1\mathbf{a}_1\right|}{\left|d\Theta^1\mathbf{a}_1\right|} = \left|K_1^1\right| \tag{1.19.25}$$

and so the mixed component K_1^1 is the curvature in the Θ^1 direction. Similarly, K_2^2 is the curvature in the Θ^2 direction.

Assume now that $K_1^1 = 0, K_1^2 \neq 0$. Eqn. 1.19.24 now reads $d\mathbf{a}_3 = -K_1^2 d\Theta^1\mathbf{a}_2$ and, referring Fig. 1.19.3, the **twist** of the surface with respect to the coordinates is

$$\frac{d\varphi}{ds} = \frac{\left|-K_1^2 d\Theta^1\mathbf{a}_2\right|}{\left|d\Theta^1\mathbf{a}_1\right|} = \left|K_1^2\right|\frac{\left|\mathbf{a}_2\right|}{\left|\mathbf{a}_1\right|} \tag{1.19.26}$$

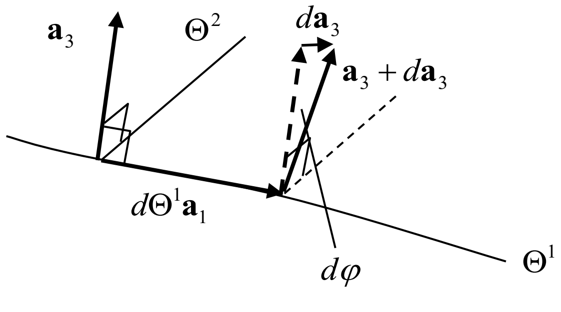


Figure 1.19.3: Twisting over the Surface

When $\left|\mathbf{a}_1\right| = \left|\mathbf{a}_2\right|$, $\left|K_1^2\right|$ is the twist; when they are not equal, $\left|K_1^2\right|$ is closely related to the twist.

[1] this is essentially the same definition as for the space curve of §1.6.2; there, the angle $\phi = \kappa\Delta s$

Two important quantities are often used to describe the curvature of a surface. These are the first and the third principal scalar invariants:

$$\mathrm{I}_{\mathbf{K}} = K^{i}_{\cdot i} = K^1_1 + K^2_2$$
$$\mathrm{III}_{\mathbf{K}} = \det K^{i}_{\cdot j} = \begin{vmatrix} K^1_1 & K^1_2 \\ K^2_1 & K^2_2 \end{vmatrix} = K^1_1 K^2_2 - K^1_2 K^2_1 = \varepsilon_{\alpha\beta} K^{\alpha}_1 K^{\beta}_2 \tag{1.19.27}$$

The first invariant is twice the **mean curvature** K_M whilst the third invariant is called the **Gaussian curvature** (or **Total curvature**) K_G of the surface.

Example (Curvature of a Sphere)

The surface of a sphere of radius *a* can be described by the coordinates (Θ^1, Θ^2), Fig. 1.19.4, where

$$x^1 = a\sin\Theta^1\cos\Theta^2, \quad x^2 = a\sin\Theta^1\sin\Theta^2, \quad x^3 = a\cos\Theta^1$$

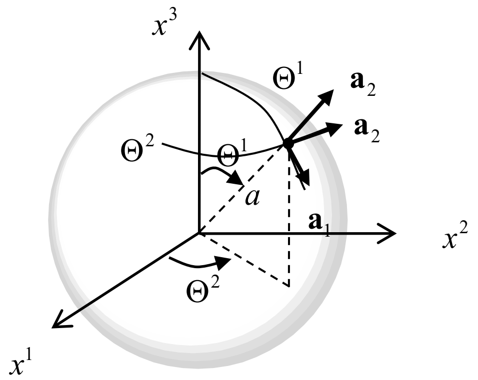


Figure 1.19.4: a spherical surface

Then, from the definitions 1.16.19, 1.16.27-28, 1.16.34, {▲Problem 2}

$$\begin{aligned}
\mathbf{a}_1 &= +a\cos\Theta^1\cos\Theta^2\mathbf{e}_1 + a\cos\Theta^1\sin\Theta^2\mathbf{e}_2 - a\sin\Theta^1\mathbf{e}_3 \\
\mathbf{a}_2 &= -a\sin\Theta^1\sin\Theta^2\mathbf{e}_1 + a\sin\Theta^1\cos\Theta^2\mathbf{e}_2 \\
\mathbf{a}^1 &= \frac{1}{a^2}\mathbf{a}_1 \\
\mathbf{a}^2 &= \frac{1}{a^2\sin^2\Theta^1}\mathbf{a}_2
\end{aligned} \tag{1.19.28}$$
$$g_{\alpha\beta} = \begin{vmatrix} a^2 & 0 \\ 0 & a^2\sin^2\Theta^1 \end{vmatrix}, \qquad g = a^4\sin^2\Theta^1$$

From 1.19.6,

$$\mathbf{a}_3 = \sin\Theta^1\cos\Theta^2\mathbf{e}_1 + \sin\Theta^1\sin\Theta^2\mathbf{e}_2 + \cos\Theta^1\mathbf{e}_3 \tag{1.19.29}$$

and this is clearly an orthogonal coordinate system with scale factors

$$h_1 = a, \quad h_2 = a\sin\Theta^1, \quad h_3 = 1 \tag{1.19.30}$$

The surface Christoffel symbols are, from 1.18.33, 1.18.36,

$$\Gamma^1_{11} = \Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{11} = \Gamma^2_{22} = 0, \quad \Gamma^2_{12} = \Gamma^2_{21} = \frac{\cos\Theta^1}{\sin^2\Theta^1}, \quad \Gamma^1_{22} = -\sin\Theta^1\cos\Theta^1 \tag{1.19.31}$$

Using the definitions 1.18.4, {▲Problem 3}

$$\begin{aligned}
&\Gamma^1_{13} = \Gamma^1_{31} = \frac{1}{a}, \quad \Gamma^1_{23} = \Gamma^1_{32} = 0 \\
&\Gamma^2_{13} = \Gamma^2_{31} = 0, \quad \Gamma^2_{23} = \Gamma^2_{32} = \frac{1}{a} \\
&\Gamma^3_{11} = -a, \quad \Gamma^3_{12} = \Gamma^3_{21} = 0, \quad \Gamma^3_{22} = -a\sin^2\Theta^1
\end{aligned} \tag{1.19.32}$$

with the remaining symbols $\Gamma^3_{\alpha 3} = \Gamma^3_{3\alpha} = \Gamma^\alpha_{33} = \Gamma^3_{33} = 0$.

The components of the curvature tensor are then, from 1.19.21, 1.19.19,

$$\left[K^\alpha_\beta\right] = \begin{bmatrix} -\frac{1}{a} & 0 \\ 0 & -\frac{1}{a} \end{bmatrix}, \quad \left[K_{\alpha\beta}\right] = \begin{bmatrix} -a & 0 \\ 0 & -a\sin^2\Theta^1 \end{bmatrix} \tag{1.19.33}$$

The mean and Gaussian curvature of a sphere are then

$$\begin{aligned}
K_M &= -\frac{2}{a} \\
K_G &= \frac{1}{a^2}
\end{aligned} \tag{1.19.34}$$

The principal curvatures are evidently K^1_1 and K^2_2. As expected, they are simply the reciprocal of the radius of curvature *a*.

■

1.19.3 Covariant Derivatives

Vectors

Consider a vector **v**, which is not necessarily a surface vector, that is, it might have a normal component $v_3 = v^3$. The covariant derivative is

$$
\begin{aligned}
v^{\alpha}\,|_{\beta} &= v^{\alpha}{}_{,\beta} + \Gamma^{\alpha}_{\gamma\beta} v^{\gamma} + \Gamma^{\alpha}_{3\beta} v^{3} & \qquad v_{\alpha}\,|_{\beta} &= v_{\alpha,\beta} - \Gamma^{\gamma}_{\alpha\beta} v_{\gamma} - \Gamma^{3}_{\alpha\beta} v_{3} \\
v^{\alpha}\,|_{3} &= v^{\alpha}{}_{,3} + \Gamma^{\alpha}_{\gamma 3} v^{\gamma} + \Gamma^{\alpha}_{33} v^{3} & v_{\alpha}\,|_{3} &= v_{\alpha,3} - \Gamma^{\gamma}_{\alpha 3} v_{\gamma} - \Gamma^{3}_{\alpha 3} v_{3} \\
&= v^{\alpha}{}_{,3} + \Gamma^{\alpha}_{\gamma 3} v^{\gamma} \qquad , & &= v_{\alpha,3} - \Gamma^{\gamma}_{\alpha 3} v_{\gamma} \\
v^{3}\,|_{\alpha} &= v^{3}{}_{,\alpha} + \Gamma^{3}_{\gamma\alpha} v^{\gamma} + \Gamma^{3}_{3\alpha} v^{3} & v_{3}\,|_{\alpha} &= v_{3,\alpha} - \Gamma^{\gamma}_{3\alpha} v_{\gamma} - \Gamma^{3}_{3\alpha} v_{3} \\
&= v^{3}{}_{,\alpha} + \Gamma^{3}_{\gamma\alpha} v^{\gamma} & &= v_{3,\alpha} - \Gamma^{\gamma}_{3\alpha} v_{\gamma}
\end{aligned}
\tag{1.19.35}
$$

Define now a two-dimensional analogue of the three-dimensional covariant derivative through

$$
\begin{aligned}
v^{\alpha}\,\|_{\beta} &= v^{\alpha}{}_{,\beta} + \Gamma^{\alpha}_{\gamma\beta} v^{\gamma} \\
v_{\alpha}\,\|_{\beta} &= v_{\alpha,\beta} - \Gamma^{\gamma}_{\alpha\beta} v_{\gamma}
\end{aligned}
\tag{1.19.36}
$$

so that, using 1.19.19, 1.19.21, the covariant derivative can be expressed as

$$
\begin{aligned}
v^{\alpha}\,|_{\beta} &= v^{\alpha}\,\|_{\beta} - K^{\alpha}_{\beta} v^{3} \\
v_{\alpha}\,|_{\beta} &= v_{\alpha}\,\|_{\beta} - K_{\alpha\beta} v_{3}
\end{aligned}
\tag{1.19.37}
$$

In the special case when the vector is a plane vector, then $v_3 = v^3 = 0$, and there is no difference between the three-dimensional and two-dimensional covariant derivatives. In the general case, the covariant derivatives can now be expressed as

$$
\begin{aligned}
\mathbf{v}_{,\beta} &= v^{i}\,|_{\beta}\, \mathbf{a}_{i} \\
&= \left(v^{\alpha}\,\|_{\beta} - K^{\alpha}_{\beta} v^{3}\right)\mathbf{a}_{\alpha} + v^{3}\,|_{\beta}\, \mathbf{a}_{3} \\
\mathbf{v}_{,\beta} &= v_{i}\,|_{\beta}\, \mathbf{a}^{i} \\
&= \left(v_{\alpha}\,\|_{\beta} - K_{\alpha\beta} v_{3}\right)\mathbf{a}^{\alpha} + v_{3}\,|_{\beta}\, \mathbf{a}^{3}
\end{aligned}
\tag{1.19.38}
$$

From 1.18.25, the gradient of a surface vector is (using 1.19.21)

$$
\operatorname{grad}\mathbf{v} = \left(v_{\alpha}\,\|_{\beta} - K_{\alpha\beta} v_{3}\right)\mathbf{a}^{\alpha} \otimes \mathbf{a}^{\beta} + K^{\gamma}_{\alpha} v_{\gamma} \mathbf{a}^{\alpha} \otimes \mathbf{a}^{3}
\tag{1.19.39}
$$

Tensors

The covariant derivatives of second order tensor components are given by 1.18.18. For example,

$$
\begin{aligned}
A^{ij}\,|_{\gamma} &= A^{ij}{}_{,\gamma} + \Gamma^{i}_{m\gamma} A^{mj} + \Gamma^{j}_{m\gamma} A^{im} \\
&= A^{ij}{}_{,\gamma} + \Gamma^{i}_{\lambda\gamma} A^{\lambda j} + \Gamma^{i}_{3\gamma} A^{3j} + \Gamma^{j}_{\lambda\gamma} A^{i\lambda} + \Gamma^{j}_{3\gamma} A^{i3}
\end{aligned}
\tag{1.19.40}
$$

Here, only surface tensors will be examined, that is, all components with an index 3 are zero. The two dimensional (plane) covariant derivative is

$$A^{\alpha\beta}\,\|_\gamma \equiv A^{\alpha\beta}{}_{,\gamma} + \Gamma^{\alpha}_{\lambda\gamma} A^{\lambda\beta} + \Gamma^{\beta}_{\lambda\gamma} A^{\alpha\lambda} \tag{1.19.41}$$

Although $A^{\alpha 3} = A^{3\alpha} = 0$ for plane tensors, one still has non-zero

$$\begin{aligned} A^{\alpha 3}\,|_\gamma &= A^{\alpha 3}{}_{,\gamma} + \Gamma^{\alpha}_{\lambda\gamma} A^{\lambda 3} + \Gamma^{3}_{\lambda\gamma} A^{\alpha\lambda} \\ &= \Gamma^{3}_{\lambda\gamma} A^{\alpha\lambda} \\ &= K_{\lambda\gamma} A^{\alpha\lambda} \\ A^{3\beta}\,|_\gamma &= A^{3\beta}{}_{,\gamma} + \Gamma^{3}_{\lambda\gamma} A^{\lambda\beta} + \Gamma^{\beta}_{\lambda\gamma} A^{3\lambda} \\ &= \Gamma^{3}_{\lambda\gamma} A^{\lambda\beta} \\ &= K_{\lambda\gamma} A^{\lambda\beta} \end{aligned} \tag{1.19.42}$$

with $A^{33}\,|_\gamma = 0$.

From 1.18.28, the divergence of a surface tensor is

$$\operatorname{div} \mathbf{A} = A^{\alpha\beta}\,\|_\beta\, \mathbf{a}_\alpha + K_{\beta\gamma} A^{\beta\gamma} \mathbf{a}_3 \tag{1.19.43}$$

1.19.4 The Gauss-Codazzi Equations

Some useful equations can be derived by considering the second derivatives of the base vectors. First, from 1.18.2,

$$\begin{aligned} \mathbf{a}_{\alpha,\beta} &= \Gamma^{\lambda}_{\alpha\beta} \mathbf{a}_\lambda + \Gamma^{3}_{\alpha\beta} \mathbf{a}_3 \\ &= \Gamma^{\lambda}_{\alpha\beta} \mathbf{a}_\lambda + K_{\alpha\beta} \mathbf{a}_3 \end{aligned} \tag{1.19.44}$$

A second derivative is

$$\mathbf{a}_{\alpha,\beta\gamma} = \Gamma^{\lambda}_{\alpha\beta,\gamma} \mathbf{a}_\lambda + \Gamma^{\lambda}_{\alpha\beta} \mathbf{a}_{\lambda,\gamma} + K_{\alpha\beta,\gamma} \mathbf{a}_3 + K_{\alpha\beta} \mathbf{a}_{3,\gamma} \tag{1.19.45}$$

Eliminating the base vectors derivatives using 1.19.44 and 1.19.20b leads to {▲Problem 4}

$$\mathbf{a}_{\alpha,\beta\gamma} = \left(\Gamma^{\lambda}_{\alpha\beta,\gamma} + \Gamma^{\eta}_{\alpha\beta}\Gamma^{\lambda}_{\eta\gamma} - K_{\alpha\beta} K^{\lambda}_{\gamma}\right)\mathbf{a}_\lambda + \left(\Gamma^{\lambda}_{\alpha\beta} K_{\lambda\gamma} + K_{\alpha\beta,\gamma}\right)\mathbf{a}_3 \tag{1.19.46}$$

This equals the partial derivative $\mathbf{a}_{\alpha,\gamma\beta}$. Comparison of the coefficient of $\mathbf{a}_3$ for these alternative expressions for the second partial derivative leads to

$$K_{\alpha\beta,\gamma} - \Gamma^{\lambda}_{\alpha\gamma} K_{\lambda\beta} = K_{\alpha\gamma,\beta} - \Gamma^{\lambda}_{\alpha\beta} K_{\lambda\gamma} \tag{1.19.47}$$

From Eqn. 1.18.18,

$$K_{\alpha\beta}\,\|_{\gamma} = K_{\alpha\beta,\gamma} - \Gamma^{\lambda}_{\alpha\gamma} K_{\lambda\beta} - \Gamma^{\lambda}_{\beta\gamma} K_{\alpha\lambda} \tag{1.19.48}$$

and so

$$K_{\alpha\beta}\,\|_{\gamma} = K_{\alpha\gamma}\,\|_{\beta} \tag{1.19.49}$$

These are the **Codazzi equations**, in which there are only two independent non-trivial relations:

$$K_{11}\,\|_{2} = K_{12}\,\|_{1}, \qquad K_{22}\,\|_{1} = K_{12}\,\|_{2} \tag{1.19.50}$$

Raising indices using the metric coefficients leads to the similar equations

$$K^{\alpha}_{\beta}\,\|_{\gamma} = K^{\alpha}_{\gamma}\,\|_{\beta} \tag{1.19.51}$$

The Riemann-Christoffel Curvature Tensor

Comparing the coefficients of $\mathbf{a}_{\lambda}$ in 1.19.46 and the similar expression for the second partial derivative shows that

$$\Gamma^{\lambda}_{\alpha\gamma,\beta} - \Gamma^{\lambda}_{\alpha\beta,\gamma} + \Gamma^{\eta}_{\alpha\gamma}\Gamma^{\lambda}_{\eta\beta} - \Gamma^{\eta}_{\alpha\beta}\Gamma^{\lambda}_{\eta\gamma} = K_{\alpha\gamma}K^{\lambda}_{\beta} - K_{\alpha\beta}K^{\lambda}_{\gamma} \tag{1.19.52}$$

The terms on the left are the two-dimensional Riemann-Christoffel, Eqn. 1.18.21, and so

$$R^{\lambda}_{\cdot\alpha\beta\gamma} = K_{\alpha\gamma}K^{\lambda}_{\beta} - K_{\alpha\beta}K^{\lambda}_{\gamma} \tag{1.19.53}$$

Further,

$$R_{\lambda\alpha\beta\gamma} = g_{\lambda\eta}R^{\eta}_{\cdot\alpha\beta\gamma} = K_{\alpha\gamma}g_{\lambda\eta}K^{\eta}_{\beta} - K_{\alpha\beta}g_{\lambda\eta}K^{\eta}_{\gamma} = K_{\alpha\gamma}K_{\beta\lambda} - K_{\alpha\beta}K_{\gamma\lambda} \tag{1.19.54}$$

These are the **Gauss equations**. From 1.18.21 *et seq.*, only 4 of the Riemann-Christoffel symbols are non-zero, and they are related through

$$R_{1212} = -R_{2112} = -R_{1221} = R_{2121} \tag{1.19.55}$$

so that there is in fact only one independent non-trivial Gauss relation. Further,

$$\begin{aligned} R_{\lambda\alpha\beta\gamma} &= K_{\alpha\gamma}K_{\beta\lambda} - K_{\alpha\beta}K_{\gamma\lambda} \\ &= K^{\mu}_{\alpha}K^{\eta}_{\lambda}\left(g_{\gamma\mu}g_{\beta\eta} - g_{\beta\mu}g_{\gamma\eta}\right) \\ &= K^{\mu}_{\alpha}K^{\eta}_{\lambda}\left(\delta^{\nu}_{\mu}\delta^{\rho}_{\eta} - \delta^{\rho}_{\mu}\delta^{\nu}_{\eta}\right)g_{\beta\rho}g_{\gamma\nu} \end{aligned} \tag{1.19.56}$$

Using 1.19.4b, 1.19.3,

$$\begin{aligned} R_{\lambda\alpha\beta\gamma} &= K_\alpha^\mu K_\lambda^\eta e^{\rho\nu} e_{\eta\mu} g_{\beta\rho} g_{\gamma\nu} \\ &= K_\alpha^\mu K_\lambda^\eta e_{\beta\gamma} e_{\eta\mu} \\ &= g\varepsilon_{\beta\gamma}\varepsilon_{\eta\mu} K_\alpha^\mu K_\lambda^\eta \end{aligned} \tag{1.19.57}$$

and so the Gauss relation can be expressed succinctly as

$$K_G = \frac{R_{1212}}{g} \tag{1.19.58}$$

where K_G is the Gaussian curvature, 1.19.27b. Thus the Riemann-Christoffel tensor is zero if and only if the Gaussian curvature is zero, and in this case only can the order of the two covariant differentiations be interchanged.

The Gauss-Codazzi equations, 1.19.50 and 1.19.58, are equivalent to a set of two first order and one second order differential equations that must be satisfied by the three independent metric coefficients $g_{\alpha\beta}$ and the three independent curvature tensor coefficients $K_{\alpha\beta}$.

Intrinsic Surface Properties

An **intrinsic** property of a surface is any quantity that remains unchanged when the surface is bent into another shape without stretching or shrinking. Some examples of intrinsic properties are the length of a curve on the surface, surface area, the components of the surface metric tensor $g_{\alpha\beta}$ (and hence the components of the Riemann-Christoffel tensor) and the Guassian curvature (which follows from the Gauss equation 1.19.58).

A **developable surface** is one which can be obtained by bending a plane, for example a piece of paper. Examples of developable surfaces are the cylindrical surface and the surface of a cone. Since the Riemann-Christoffel tensor and hence the Gaussian curvature vanish for the plane, they vanish for all developable surfaces.

1.19.5 Geodesics

The Geodesic Curvature and Normal Curvature

Consider a curve C lying on the surface, with arc length s measured from some fixed point. As for the space curve, §1.6.2, one can define the unit tangent vector $\boldsymbol{\tau}$, principal normal $\mathbf{v}$ and binormal vector $\mathbf{b}$ (Eqn. 1.6.3 *et seq.*):

$$\boldsymbol{\tau} = \frac{d\mathbf{x}}{ds} = \frac{d\Theta^\alpha}{ds}\mathbf{a}_\alpha, \qquad \mathbf{v} = \frac{1}{\kappa}\frac{d\boldsymbol{\tau}}{ds}, \qquad \mathbf{b} = \boldsymbol{\tau}\times\mathbf{v} \tag{1.19.59}$$

so that the curve passes along the intersection of the osculating plane containing $\boldsymbol{\tau}$ and $\mathbf{v}$ (see Fig. 1.6.3), and the surface These vectors form an orthonormal set but, although $\mathbf{v}$ is normal to the tangent, it is not necessarily normal to the surface, as illustrated in Fig.

1.19.5. For this reason, form the new orthonormal triad $(\boldsymbol{\tau}, \boldsymbol{\tau}_2, \mathbf{a}_3)$, so that the unit vector $\boldsymbol{\tau}_2$ lies in the plane tangent to the surface. From 1.19.59, 1.19.3,

$$\boldsymbol{\tau}_2 = \mathbf{a}_3 \times \boldsymbol{\tau} = \frac{d\Theta^\alpha}{ds}\mathbf{a}_3 \times \mathbf{a}_\alpha = e_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\mathbf{a}^\beta \tag{1.19.60}$$

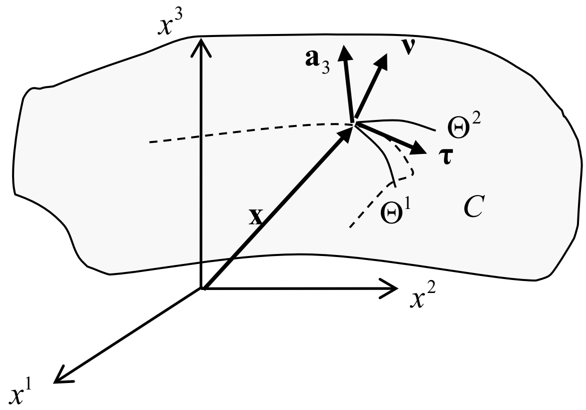


Figure 1.19.5: a curve lying on a surface

Next, the vector $d\boldsymbol{\tau}/ds$ will be decomposed into components along $\boldsymbol{\tau}_2$ and the normal $\mathbf{a}_3$. First, differentiate 1.19.59a and use 1.19.44b to get {▲Problem 5}

$$\frac{d\boldsymbol{\tau}}{ds} = \left(\frac{d^2\Theta^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\right)\mathbf{a}_\gamma + K_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\mathbf{a}_3 \tag{1.19.61}$$

Then

$$\frac{d\boldsymbol{\tau}}{ds} = \kappa_g \boldsymbol{\tau}_2 + \kappa_n \mathbf{a}_3 \tag{1.19.62}$$

where

$$\begin{aligned}\kappa_g &= e_{\lambda\gamma}\frac{d\Theta^\lambda}{ds}\left(\frac{d^2\Theta^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\right) \\ \kappa_n &= K_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\end{aligned} \tag{1.19.63}$$

These are formulae for the **geodesic curvature** κ_g and the **normal curvature** κ_n. Many different curves with representations $\Theta^\alpha(s)$ can pass through a certain point with a given tangent vector $\boldsymbol{\tau}$. Form 1.19.59, these will all have the same value of $d\Theta^\alpha / ds$ and so, from 1.19.63, these curves will have the same normal curvature but, in general, different geodesic curvatures.

A curve passing through a **normal section**, that is, along the intersection of a plane containing $\boldsymbol{\tau}$ and $\mathbf{a}_3$, and the surface, will have zero geodesic curvature.

The normal curvature can be expressed as

$$\kappa_n = \boldsymbol{\tau}\, \mathbf{K}\, \boldsymbol{\tau} \tag{1.19.64}$$

If the tangent is along an eigenvector of **K**, then κ_n is an eigenvalue, and hence a maximum or minimum normal curvature. Surface curves with the property that an eigenvector of the curvature tensor is tangent to it at every point is called a **line of curvature**. A convenient coordinate system for a surface is one in which the coordinate curves are lines of curvature. Such a system, with Θ^1 containing the maximum values of κ_n, has at every point a curvature tensor of the form

$$\left[K_i^{\,j}\right] = \begin{bmatrix} K_1^1 & 0 \\ 0 & K_2^2 \end{bmatrix} = \begin{bmatrix} (\kappa_n)_{\max} & 0 \\ 0 & (\kappa_n)_{\min} \end{bmatrix} \tag{1.19.65}$$

This was the case with the spherical surface example discussed in §1.19.2.

The Geodesic

A **geodesic** is defined to be a curve which has zero geodesic curvature *at every point* along the curve. Form 1.19.63, parametric equations for the geodesics over a surface are

$$\frac{d^2\Theta^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds} = 0 \tag{1.19.64}$$

It can be proved that the geodesic is the curve of shortest distance joining two points on the surface. Thus the geodesic curvature is a measure of the deviance of the curve from the shortest-path curve.

The Geodesic Coordinate System

If the Gaussian curvature of a surface is not zero, then it is not possible to find a surface coordinate system for which the metric tensor components $g_{\alpha\beta}$ equal the Kronecker delta $\delta_{\alpha\beta}$ everywhere. Such a geometry is called **Riemannian**. However, it is always possible to construct a coordinate system in which $g_{\alpha\beta} = \delta_{\alpha\beta}$, and the derivatives of the metric coefficients are zero, *at a particular point* on the surface. This is the **geodesic coordinate system**.

1.19.6 Problems

1 Derive Eqns. 1.19.16, $\Gamma^3_{\alpha 3} = \Gamma^\alpha_{33} = \Gamma^3_{33} = 0$.

2 Derive the Cartesian components of the curvilinear base vectors for the spherical surface, Eqn. 1.19.28.

3 Derive the Christoffel symbols for the spherical surface, Eqn. 1.19.32.
4 Use Eqns. 1.19.44-5 and 1.19.20b to derive 1.19.46.
5 Use Eqns. 1.19.59a and 1.19.44b to derive 1.19.61.

1.A Appendix to Chapter 1

1.A.1 The Algebraic Structures of Groups, Fields and Rings

Definition:
The nonempty set *G* with a binary operation, that is, to each pair of elements $a, b \in G$ there is assigned an element $ab \in G$, is called a **group** if the following axioms hold:

1. *associative law*: $(ab)c = a(bc)$ for any $a, b, c \in G$
2. *identity element*: there exists an element $e \in G$, called the identity element, such that $ae = ea = a$
3. *inverse*: for each $a \in G$, there exists an element $a^{-1} \in G$, called the inverse of *a*, such that $aa^{-1} = a^{-1}a = e$

Examples:

(a) An example of a group is the set of integers under addition. In this case the binary operation is denoted by +, as in $a + b$; one has (1) addition is associative, $(a+b)+c$ equals $a+(b+c)$, (2) the identity element is denoted by 0, $a + 0 = 0 + a = a$, (3) the inverse of *a* is denoted by *a*, called the *negative* of *a*, and $a + (-a) = (-a) + a = 0$

Definition:
An **abelian group** is one for which the commutative law holds, that is, if $ab = ba$ for every $a, b \in G$.

Examples:

(a) The above group, the set of integers under addition, is commutative, $a + b = b + a$, and so is an abelian group.

Definition:
A mapping *f* of a group *G* to another group G', $f : G \to G'$, is called a **homomorphism** if $f(ab) = f(a)f(b)$ for every $a, b \in G$; if *f* is bijective (one-one and onto), then it is called an **isomorphism** and *G* and G' re said to be **isomorphic**

Definition:
If $f : G \to G'$ is a homomorphism, then the **kernel** of *f* is the set of elements of *G* which map into the identity element of G', $k = \{a \in G \mid f(a) = e'\}$

Examples

(a) Let *G* be the group of non-zero complex numbers under multiplication, and let G' be the non-zero real numbers under multiplication. The mapping $f : G \to G'$ defined by $f(z) = |z|$ is a homomorphism, because

$$f(z_1 z_2) = |z_1 z_2| = |z_1||z_2| = f(z_1) f(z_2)$$

The kernel of *f* is the set of elements which map into 1, that is, the complex numbers on the unit circle

Definition:
The non-empty set *A* with the two binary operations of addition (denoted by +) and multiplication (denoted by juxtaposition) is called a **ring** if the following are satisfied:

1. *associative law for addition*: for any $a, b, c \in A$, $(a+b)+c = a+(b+c)$
2. *zero element* (additive identity): there exists an element $0 \in A$, called the zero element, such that $a+0 = 0+a = a$ for every $a \in A$
3. *negative* (additive inverse): for each $a \in A$ there exists an element $-a \in A$, called the negative of *a*, such that $a+(-a) = (-a)+a = 0$
4. *commutative law for addition*: for any $a, b \in A$, $a+b = b+a$
5. *associative law for multiplication*: for any $a, b, c \in A$, $(ab)c = a(bc)$
6. *distributive law of multiplication over addition* (both left and right distributive): for any $a, b, c \in A$, (i) $a(b+c) = ab+ac$, (ii) $(b+c)a = ba+ca$

Remarks:
(i) the axioms 1-4 may be summarized by saying that *A* is an abelian group under addition
(ii) the operation of **subtraction** in a ring is defined through $a-b \equiv a+(-b)$
(iii) using these axioms, it can be shown that $a0 = 0a = 0$, $a(-b) = (-a)b = -ab$, $(-a)(-b) = ab$ for all $a, b \in A$

Definition:
A **commutative ring** is a ring with the additional property:

7. *commutative law for multiplication*: for any $a, b \in A$, $ab = ba$

Definition:
A **ring with a unit element** is a ring with the additional property:

8. *unit element* (multiplicative identity): there exists a nonzero element $1 \in A$ such that $a1 = 1a = a$ for every $a \in A$

Definition:
A commutative ring with a unit element is an **integral domain** if it has no zero divisors, that is, if $ab = 0$, then $a = 0$ or $b = 0$

Examples:
(a) the set of integers *Z* is an integral domain

Definition:
A commutative ring with a unit element is a **field** if it has the additional property:

9. *multiplicative inverse*: there exists an element $a^{-1} \in A$ such that $aa^{-1} = a^{-1}a = 1$

Remarks:
(i) note that the number 0 has no multiplicative inverse. When constructing the real numbers *R*, 0 is a special element which is not allowed have a multiplicative inverse. For this reason, division by 0 in *R* is indeterminate

Examples:

(a) The set of real numbers R with the usual operations of addition and multiplication forms a field

(b) The set of ordered pairs of real numbers with addition and multiplication defined by

$$(a,b)+(c,d)=(a+c,b+d)$$
$$(a,b)(c,d)=(ac-bd,ad+bc)$$

is also a field - this is just the set of complex numbers C

1.A.2 The Linear (Vector) Space

Definition:

Let F be a given field whose elements are called *scalars*. Let V be a non-empty set with rules of addition and scalar multiplication, that is there is a *sum* $a+b$ for any $a, b \in V$ and a *product* αa for any $a \in V$, $\alpha \in F$. Then V is called a **linear space** over F if the following eight axioms hold:

1. *associative law for addition*: for any $a, b, c \in V$, one has $(a+b)+c = a+(b+c)$
2. *zero element*: there exists an element $0 \in V$, called the zero element, or origin, such that $a+0=0+a=a$ for every $a \in V$
3. *negative*: for each $a \in V$ there exists an element $-a \in V$, called the negative of *a*, such that $a+(-a)=(-a)+a=0$
4. *commutative law for addition*: for any $a, b \in V$, we have $a+b=b+a$
5. *distributive law, over addition of elements of V*: for any $a, b \in V$ and scalar $\alpha \in F$, $\alpha(a+b)=\alpha a+\alpha b$
6. *distributive law, over addition of scalars*: for any $a \in V$ and scalars $\alpha, \beta \in F$, $(\alpha+\beta)a=\alpha a+\beta a$
7. *associative law for multiplication*: for any $a \in V$ and scalars $\alpha, \beta \in F$, $\alpha(\beta a)=(\alpha\beta)a$
8. *unit multiplication*: for the unit scalar $1 \in F$, $1a = a$ for any $a \in V$.

1.B Appendix to Chapter 1

1.B.1 The Ordinary Calculus

Here are listed some important concepts from the ordinary calculus.

The Derivative

Consider *u*, a function *f* of one independent variable *x*. The *derivative* of *u* at *x* is defined by

$$\frac{du}{dx} \equiv f'(x) = \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{1.B.1}$$

where Δu is the **increment** in *u* due to an increment Δx in *x*.

The Differential

The **differential** of *u* is defined by

$$du = f'(x)\Delta x \tag{1.B.2}$$

By considering the special case of $u = f(x) = x$, one has $du = dx = \Delta x$, so the differential of the independent variable is equivalent to the increment. $dx = \Delta x$. Thus, in general, the differential can be written as $du = f'(x)dx$. The differential of *u* and increment in *u* are only approximately equal, $du \approx \Delta u$, and approach one another as $\Delta x \to 0$. This is illustrated in Fig. 1.B.1.

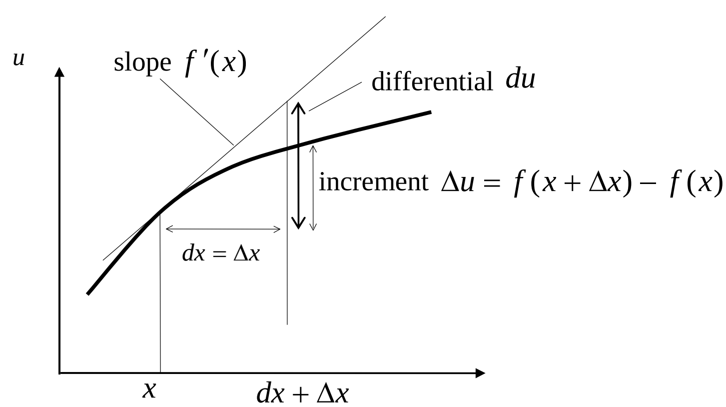


Figure 1.B.1: the differential

If *x* is itself a function of another variable, *t* say, $u(x(t))$, then the **chain rule** of differentiation gives

$$\frac{du}{dt} = f'(x)\frac{dx}{dt} \tag{1.B.3}$$

Arc Length

The length of an arc, measured from a fixed point *a* on the arc, to *x*, is, from the definition of the integral,

$$s = \int_a^x ds = \int_a^x \sec\psi dx = \int_a^x \sqrt{1+(dy/dx)^2}\,dx \tag{1.B.4}$$

where ψ is the angle the tangent to the arc makes with the *x* axis, Fig 1.B.2, with $(dy/dx) = \tan\psi$ and $(ds)^2 = (dx)^2 + (dy)^2$ (ds is the length of the dotted line in Fig. 1.B.2b). Also, it can be seen that

$$\begin{aligned} \lim_{\Delta s \to 0} \frac{|pq|_{\text{chord}}}{|pq|_{\text{arc}}} &= \lim_{\Delta s \to 0} \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\Delta s} \\ &= \lim_{\Delta s \to 0} \sqrt{\left(\frac{\Delta x}{\Delta s}\right)^2 + \left(\frac{\Delta y}{\Delta s}\right)^2} \\ &= \sqrt{\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2} = 1 \end{aligned} \tag{1.B.5}$$

so that, if the increment Δs is small, $(\Delta s)^2 \approx (\Delta x)^2 + (\Delta y)^2$.

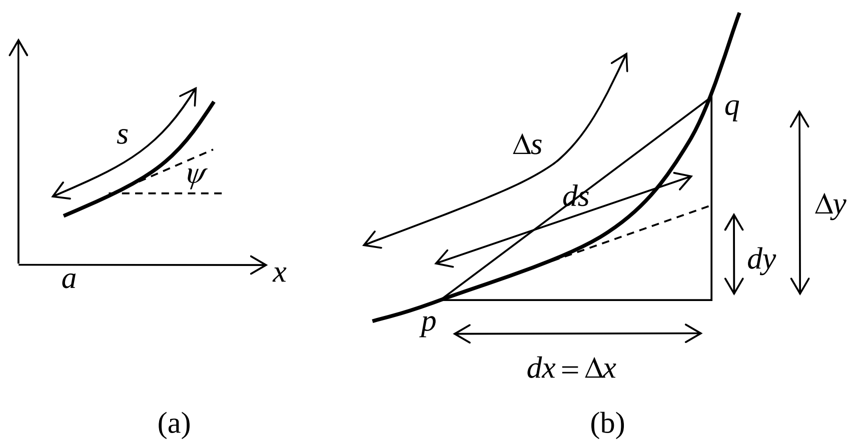


Figure 1.B.2: arc length

The Calculus of Two or More Variables

Consider now two independent variables, $u = f(x, y)$. We can define **partial derivatives** so that, for example,

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \left. \frac{\Delta u}{\Delta x} \right|_{y \text{ constant}} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \tag{1.B.6}$$

The **total differential** du due to increments in both *x* and *y* can in this case be shown to be

$$du = \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y \tag{1.B.7}$$

which is written as $du = (\partial u / \partial x)dx + (\partial u / \partial y)dy$, by setting $dx = \Delta x, dy = \Delta y$. Again, the differential du is only an approximation to the actual increment Δu (the increment and differential are shown in Fig. 1.B.3 for the case $dy = \Delta y = 0$).

It can be shown that this expression for the differential du holds whether *x* and *y* are independent, or whether they are functions themselves of an independent variable *t*, $u \equiv u(x(t), y(t))$, in which case one has the total derivative of *u* with respect to *t*,

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} \tag{1.B.8}$$

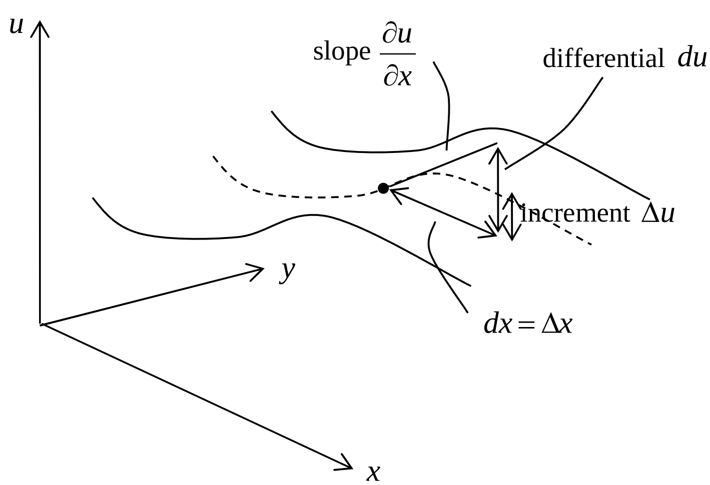


Figure 1.B.3: the partial derivative

The Chain rule for Two or More Variables

Consider the case where *u* is a function of the two variables x, y, $u = f(x, y)$, but also that *x* and *y* are functions of the two independent variables *s* and *t*, $u = f(x(s,t), y(s,t))$. Then

$$\begin{aligned} du &= \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy \\ &= \frac{\partial u}{\partial x}\left(\frac{\partial x}{\partial s}ds + \frac{\partial x}{\partial t}dt\right) + \frac{\partial u}{\partial y}\left(\frac{\partial y}{\partial s}ds + \frac{\partial y}{\partial t}dt\right) \\ &= \left(\frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s}\right)ds + \left(\frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t}\right)dt \end{aligned} \tag{1.B.9}$$

But, also,

$$du = \frac{\partial u}{\partial s}ds + \frac{\partial u}{\partial t}dt \tag{1.B.10}$$

Comparing the two, and since ds, dt are independent and arbitrary, one obtains the chain rule

$$\begin{aligned} \frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial u}{\partial t} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} \end{aligned} \tag{1.B.11}$$

In the special case when *x* and *y* are functions of only *one* variable, *t* say, so that $u = f\left[x(t), y(t)\right]$, the above reduces to the total derivative given earlier.

One can further specialise: In the case when *u* is a function of *x* and *t*, with $x = x(t)$, $u = f\left[x(t), t\right]$, one has

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial t} \tag{1.B.12}$$

When *u* is a function of one variable only, *x* say, so that $u = f\left[x(t)\right]$, the above reduces to the chain rule for ordinary differentiation.

Taylor's Theorem

Suppose the value of a function $f(x, y)$ is known at (x_0, y_0). Its value at a neighbouring point $(x_0 + \Delta x, y_0 + \Delta y)$ is then given by

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) &+ \left(\Delta x \left.\frac{\partial f}{\partial x}\right|_{(x_0, y_0)} + \Delta y \left.\frac{\partial f}{\partial y}\right|_{(x_0, y_0)}\right) \\ &+ \frac{1}{2}\left((\Delta x)^2 \left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0, y_0)} + \Delta x \Delta y \left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0, y_0)} + (\Delta x)^2 \left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0, y_0)}\right) + \cdots \end{aligned} \tag{1.B.13}$$

The Mean Value Theorem

If $f(x)$ is continuous over an interval $a < x < b$, then

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \tag{1.B.14}$$

Geometrically, this is equivalent to saying that there exists at least one point in the interval for which the tangent line is parallel to the line joining $f(a)$ and $f(b)$. This result is known as the **mean value theorem**.

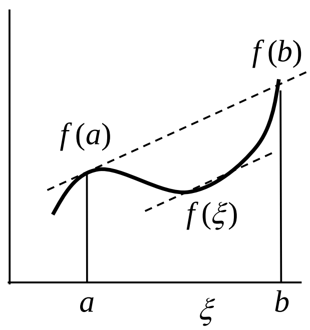


Figure 1.B.4: the mean value theorem

The law of the mean can also be written in terms of an integral: there is at least one point ξ in the interval $[a,b]$ such that

$$f(\xi) = \frac{1}{l}\int_a^b f(x)dx \tag{1.B.15}$$

where *l* is the length of the interval, $l = b - a$. The right hand side here can be interpreted as the average value of *f* over the interval. The theorem therefore states that the average value of the function lies somewhere in the interval. The equivalent expression for a double integral is that there is at least one point (ξ_1, ξ_2) in a region *R* such that

$$f(\xi_1, \xi_2) = \frac{1}{A}\iint_R f(x_1, x_2)dx_1 dx_2 \tag{1.B.16}$$

where *A* is the area of the region of integration *R*, and similarly for a triple/volume integral.

1.B.2 Transformation of Coordinate System

Let the coordinates of a point in space be (x_1, x_2, x_3). Introduce a second set of coordinates $(\Theta_1, \Theta_2, \Theta_3)$, related to the first set through the transformation equations

$$\Theta_i = f_i(x_1, x_2, x_3) \tag{1.B.17}$$

with the inverse equations

$$x_i = g_i(\Theta_1, \Theta_2, \Theta_3) \tag{1.B.18}$$

A transformation is termed an **admissible transformation** if the inverse transformation exists and is in one-to-one correspondence in a certain region of the variables (x_1, x_2, x_3), that is, each set of numbers $(\Theta_1, \Theta_2, \Theta_3)$ defines a unique set (x_1, x_2, x_3) in the region, and *vice versa*.

Now suppose that one has a point with coordinates x_i^0, Θ_i^0 which satisfy 1.B.17. Eqn. 1.B.17 will be in general non-linear, but differentiating leads to

$$d\Theta_i = \frac{\partial f_i}{\partial x_j} dx_j, \tag{1.B.19}$$

which is a system of three linear equations. From basic linear algebra, this system can be solved for the dx_j if and only if the determinant of the coefficients does not vanish, i.e

$$J = \det\left[\frac{\partial f_i}{\partial x_j}\right] \neq 0, \tag{1.B.20}$$

with the partial derivatives evaluated at x_i^0 (the one dimensional situation is shown in Fig. 1.B.5). If $J \neq 0$, one can solve for the dx_i:

$$dx_i = A_{ij} d\Theta_j, \tag{1.B.21}$$

say. This is a linear approximation of the inverse equations 1.B.18 and so the inverse exists in a small region near (x_1^0, x_2^0, x_3^0). This argument can be extended to other neighbouring points and the region in which $J \neq 0$ will be the region for which the transformation will be admissible.

If the Jacobian is positive everywhere, then a right handed set will be transformed into another right handed set, and the transformation is said to be **proper**.

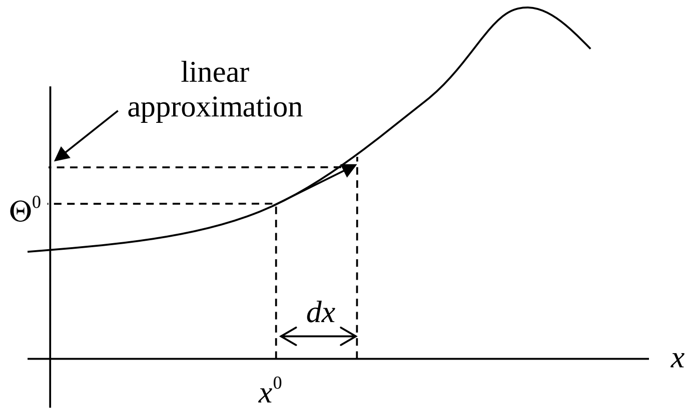


Figure 1.B.5: linear approximation

2 Kinematics

Kinematics is concerned with expressing in mathematical form the deformation and motion of materials. In what follows, a number of important quantities, mainly vectors and second-order tensors, are introduced. Each of these quantities, for example the velocity, deformation gradient or rate of deformation tensor, allows one to describe a particular aspect of a deforming material.

No consideration is given to what is *causing* the deformation and movement – the cause is the action of forces on the material, and these will be discussed in the next chapter.

The first section introduces the material and spatial coordinates and descriptions. The second and third sections discuss the strain tensors. The fourth, fifth and sixth sections deal with rates of deformation and rates of change of kinematic quantities. The theory is specialised to small strain deformations in section 7. The notion of objectivity and the related topic of rigid rotations are discussed in sections 8 and 9. The final sections, 10-13, deal with kinematics using the convected coordinate system, and include the important notions of push-forward, pull-back and the Lie time derivative.

2.1 Motion

2.1.1 The Material Body and Motion

Physical materials in the real world are modeled using an abstract mathematical entity called a **body**. This body consists of an infinite number of **material particles**[1]. Shown in Fig. 2.1.1a is a body *B* with material particle *P*. One distinguishes between this body and the space in which it resides and through which it travels. Shown in Fig. 2.1.1b is a certain **point** $\mathbf{x}$ in Euclidean point space *E*.

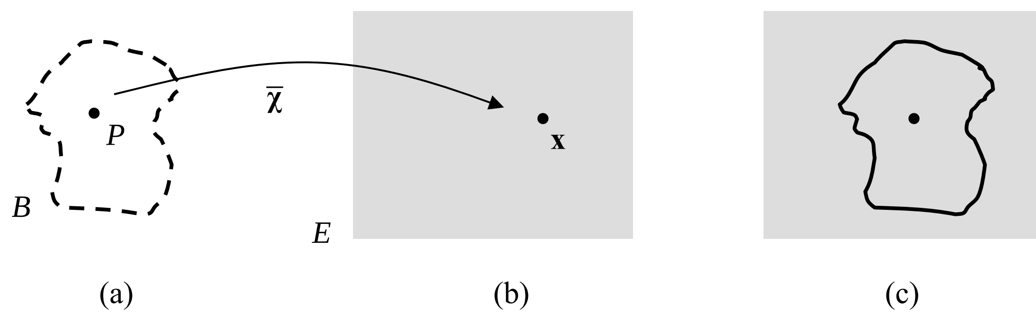


Figure 2.1.1: (a) a material particle in a body, (b) a place in space, (c) a configuration of the body

By fixing the material particles of the body to points in space, one has a **configuration** of the body $\bar{\chi}$, Fig. 2.1.1c. A configuration can be expressed as a mapping of the particles P to the point $\mathbf{x}$,

$$\mathbf{x} = \bar{\chi}(P) \tag{2.1.1}$$

A **motion** of the body is a *family* of configurations parameterised by time *t,*

$$\mathbf{x} = \bar{\chi}(P, t) \tag{2.1.2}$$

At any time *t*, Eqn. 2.1.2 gives the location in space $\mathbf{x}$ of the material particle P, Fig. 2.1.2.

[1] these particles are not the discrete mass particles of Newtonian mechanics, rather they are very small portions of continuous matter; the meaning of particle is made precise in the definitions which follow

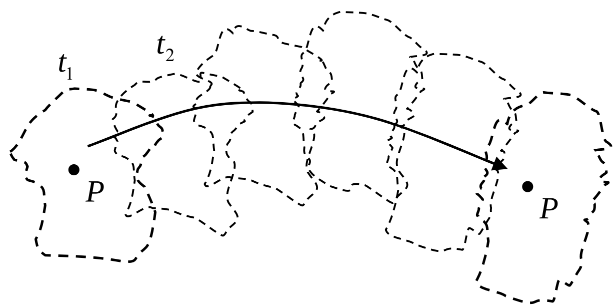


Figure 2.1.2: a motion of material

The Reference and Current Configurations

Choose now some **reference configuration**, Fig. 2.1.3. The motion can then be measured relative to this configuration. The reference configuration might be the configuration occupied by the material at time $t = 0$, in which case it is often called the **initial configuration**. For a solid, it might be natural to choose a configuration for which the material is stress-free, in which case it is often called the **undeformed configuration**. However, the choice of reference configuration is completely arbitrary.

Introduce a Cartesian coordinate system with base vectors $\mathbf{E}_i$ for the reference configuration. A material particle P in the reference configuration can then be assigned a unique position vector $\mathbf{X} = X_i\mathbf{E}_i$ relative to the origin of the axes. The coordinates (X_1, X_2, X_3) of the particle are called **material coordinates** (or **Lagrangian coordinates** or **referential coordinates**).

Some time later, say at time *t*, the material occupies a different configuration, which will be called the **current configuration** (or **deformed configuration**). Introduce a second Cartesian coordinate system with base vectors $\mathbf{e}_i$ for the current configuration, Fig. 2.1.3. In the current configuration, the same particle P now occupies the location $\mathbf{x}$, which can now also be assigned a position vector $\mathbf{x} = x_i\mathbf{e}_i$. The coordinates (x_1, x_2, x_3) are called **spatial coordinates** (or **Eulerian coordinates**).

Each particle thus has two sets of coordinates associated with it. The particle's material coordinates stay with it throughout its motion. The particle's spatial coordinates change as it moves.

The coordinate systems do not have to be Cartesian. For example, suppose one has a rectangular block which deforms into a curved beam (part of a circle). In that case it would be sensible to employ a rectangular Cartesian coordinate system with coordinates (X_1, X_2, X_3) to describe the reference configuration, and a polar coordinate system (r, θ, z) to describe the current configuration.

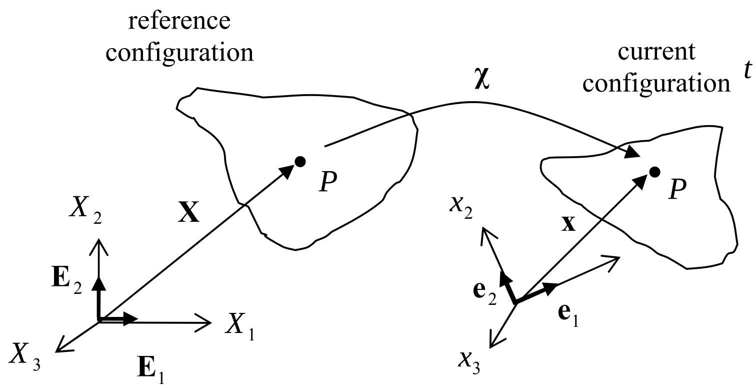


Figure 2.1.3: reference and current configurations

In practice, the material and spatial axes are usually taken to be coincident so that the base vectors $\mathbf{E}_i$ and $\mathbf{e}_i$ are the same, as in Fig. 2.1.4. Nevertheless, the use of different base vectors $\mathbf{E}$ and $\mathbf{e}$ for the reference and current configurations is useful even when the material and spatial axes are coincident, since it helps distinguish between quantities associated with the reference configuration and those associated with the spatial configuration (see later).

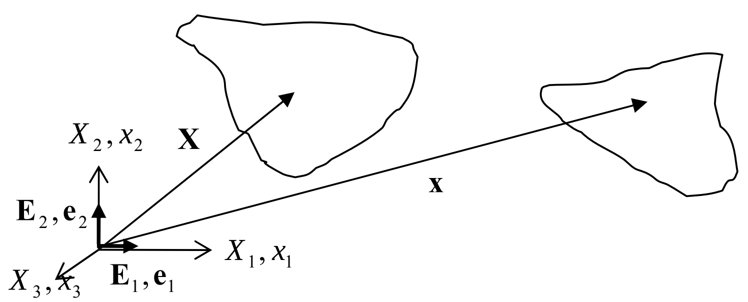


Figure 2.1.4: reference and current configurations with coincident axes

In terms of the position vectors, the motion 2.1.2 can be expressed as a relationship between the material and spatial coordinates,

$$\boxed{\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t), \qquad x_i = \chi_i(X_1, X_2, X_3, t)} \qquad \textbf{Material description} \qquad (2.1.3)$$

or the inverse relation

$$\boxed{\mathbf{X} = \boldsymbol{\chi}^{-1}(\mathbf{x}, t), \qquad X_i = \chi_i^{-1}(x_1, x_2, x_3, t)} \qquad \textbf{Spatial description} \qquad (2.1.4)$$

If one knows the material coordinates of a particle then its position in the current configuration can be determined from 2.1.3. Alternatively, if one focuses on some location in space, in the current configuration, then the material particle occupying that position can be determined from 2.1.4. This is illustrated in the following example.

Example (Extension of a Bar)

Consider the motion

$$x_1 = 3X_1 t + X_1 + t, \quad x_2 = X_2, \quad x_3 = X_3 \tag{2.1.5}$$

These equations are of the form 2.1.3 and say that "the particle that was originally at position **X** is now, at time *t*, at position **x**". They represent a simple translation and uniaxial extension of material as shown in Fig. 2.1.5. Note that $\mathbf{X} = \mathbf{x}$ at $t = 0$.

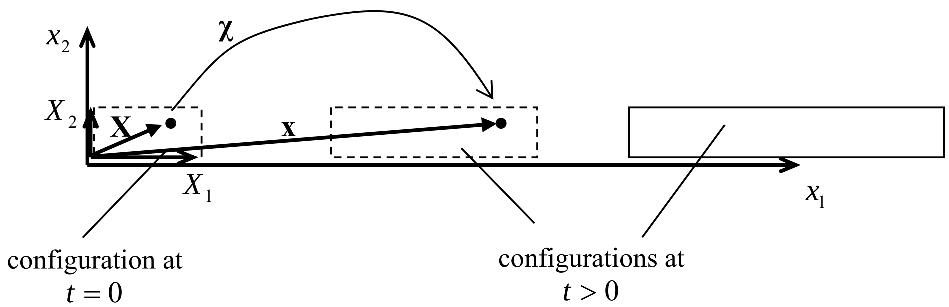


Figure 2.1.5: translation and extension of material

Relations of the form 2.1.4 can be obtained by inverting 2.1.5:

$$X_1 = \frac{x_1 - t}{1 + 3t}, \quad X_2 = x_2, \quad X_3 = x_3$$

These equations say that "the particle that is now, at time *t*, at position **x** was originally at position **X**".

■

Convected Coordinates

The material and spatial coordinate systems used here are fixed Cartesian systems. An alternative method of describing a motion is to *attach* the material coordinate system to the material and let it deform with the material. The motion is then described by defining how this coordinate system changes. This is the **convected coordinate system**. In general, the axes of a convected system will not remain mutually orthogonal and a curvilinear system is required. Convected coordinates will be examined in §2.10.

2.1.2 The Material and Spatial Descriptions

Any physical property (such as density, temperature, etc.) or kinematic property (such as displacement or velocity) of a body can be described in terms of either the material coordinates **X** or the spatial coordinates **x**, since they can be transformed into each other using 2.1.3-4. A **material** (or **Lagrangian**) **description** of events is one where the material coordinates are the independent variables. A **spatial** (or **Eulerian**) description of events is one where the spatial coordinates are used.

Example (Temperature of a Body)

Suppose the temperature θ of a body is, in material coordinates,

$$\theta(\mathbf{X},t) = 3X_1 - X_3 \tag{2.1.6}$$

but, in the spatial description,

$$\theta(\mathbf{x},t) = \frac{x_1}{t} - 1 - x_3 . \tag{2.1.7}$$

According to the material description 2.1.6, the temperature is different for different particles, but the temperature of each particle remains constant over time. The spatial description 2.1.7 describes the time-dependent temperature at a specific location in space, **x**, Fig. 2.1.6. Different material particles are flowing through this location over time.

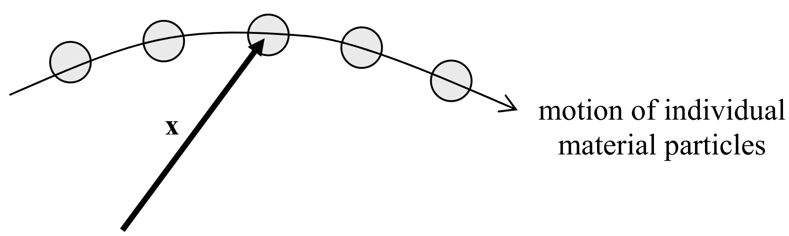


Figure 2.1.6: particles flowing through space

■

In the material description, then, attention is focused on specific *material*. The piece of matter under consideration may change shape, density, velocity, and so on, but it is always the same piece of material. On the other hand, in the spatial description, attention is focused on a fixed location in *space*. Material may pass through this location during the motion, so different material is under consideration at different times.

The spatial description is the one most often used in Fluid Mechanics since there is no natural reference configuration of the material as it is continuously moving. However, both the material and spatial descriptions are used in Solid Mechanics, where the reference configuration is usually the stress-free configuration.

2.1.3 Small Perturbations

A large number of important problems involve materials which deform only by a relatively small amount. An example would be the steel structural columns in a building under modest loading. In this type of problem there is virtually no distinction to be made between the two viewpoints taken above and the analysis is simplified greatly (see later, on Small Strain Theory, §2.7).

2.1.4 Problems

1. The density of a material is given by $\rho = 3X_1 + X_2$ and the motion is given by the equations $X_1 = x_1,\, X_2 = x_2 - t,\, X_3 = x_3 - t$.
   (a) what kind of description is this for the density, and what kind of description is this for the motion?
   (b) re-write the density in terms of **x** – what is the name given to this description of the density?
   (c) is the density of any given material particle changing with time?
   (d) invert the motion equations so that **X** is the independent variable – what is the name given to this description of the motion?
   (e) draw the line element joining the origin to $(1,1,0)$ and sketch the position of this element of material at times $t = 1$ and $t = 2$.

2.2 Deformation and Strain

A number of useful ways of describing and quantifying the deformation of a material are discussed in this section.

Attention is restricted to the reference and current configurations. No consideration is given to the particular sequence by which the current configuration is reached from the reference configuration and so the deformation can be considered to be independent of time. In what follows, particles in the reference configuration will often be termed "undeformed" and those in the current configuration "deformed".

In a change from Chapter 1, lower case letters will now be reserved for both vector- *and* tensor- functions of the spatial coordinates **x**, whereas upper-case letters will be reserved for functions of material coordinates **X**. There will be exceptions to this, but it should be clear from the context what is implied.

2.2.1 The Deformation Gradient

The **deformation gradient F** is the fundamental measure of deformation in continuum mechanics. It is the second order tensor which maps line elements in the reference configuration into line elements (consisting of the *same* material particles) in the current configuration.

Consider a line element $d\mathbf{X}$ emanating from position **X** in the reference configuration which becomes $d\mathbf{x}$ in the current configuration, Fig. 2.2.1. Then, using 2.1.3,

$$\begin{aligned} d\mathbf{x} &= \boldsymbol{\chi}(\mathbf{X} + d\mathbf{X}) - \boldsymbol{\chi}(\mathbf{X}) \\ &= (\mathrm{Grad}\,\boldsymbol{\chi})d\mathbf{X} \end{aligned} \tag{2.2.1}$$

A capital G is used on "Grad" to emphasise that this is a gradient with respect to the material coordinates[1], the **material gradient**, $\partial\boldsymbol{\chi}/\partial\mathbf{X}$.

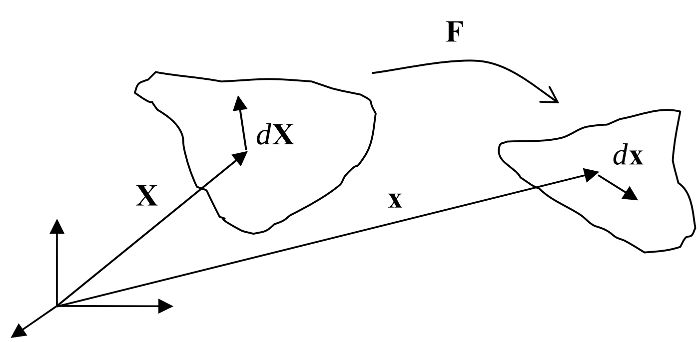


Figure 2.2.1: the Deformation Gradient acting on a line element

[1] one can have material gradients and spatial gradients of material or spatial fields – see later

The motion vector-function $\boldsymbol{\chi}$ in 2.1.3, 2.2.1, is a function of the variable **X**, but it is customary to denote this simply by **x**, the value of $\boldsymbol{\chi}$ at **X**, i.e. $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$, so that

$$\boxed{\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \operatorname{Grad}\mathbf{x}, \quad F_{iJ} = \frac{\partial x_i}{\partial X_J}} \quad \textbf{Deformation Gradient} \tag{2.2.2}$$

with

$$\boxed{d\mathbf{x} = \mathbf{F}\, d\mathbf{X}, \qquad dx_i = F_{iJ}\, dX_J} \quad \textbf{action of F} \tag{2.2.3}$$

Lower case indices are used in the index notation to denote quantities associated with the spatial basis $\{\mathbf{e}_i\}$ whereas upper case indices are used for quantities associated with the material basis $\{\mathbf{E}_I\}$.

Note that

$$d\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} d\mathbf{X}$$

is a differential quantity and this expression has some error associated with it; the error (due to terms of order $(d\mathbf{X})^2$ and higher, neglected from a Taylor series) tends to zero as the differential $d\mathbf{X} \to 0$. The deformation gradient (whose components are finite) thus characterises the deformation in the *neighbourhood* of a point **X**, mapping infinitesimal line elements $d\mathbf{X}$ emanating from **X** in the reference configuration to the infinitesimal line elements $d\mathbf{x}$ emanating from **x** in the current configuration, Fig. 2.2.2.

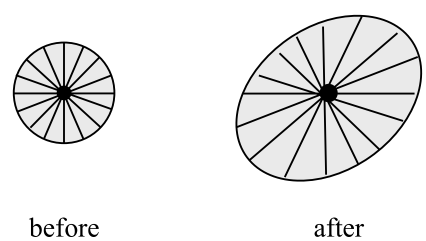


Figure 2.2.2: deformation of a material particle

Example

Consider the cube of material with sides of unit length illustrated by dotted lines in Fig. 2.2.3. It is deformed into the rectangular prism illustrated (this could be achieved, for example, by a continuous rotation and stretching motion). The material and spatial coordinate axes are coincident. The material description of the deformation is

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}) = -6X_2\mathbf{e}_1 + \frac{1}{2}X_1\mathbf{e}_2 + \frac{1}{3}X_3\mathbf{e}_3$$

and the spatial description is

$$\mathbf{X} = \boldsymbol{\chi}^{-1}(\mathbf{x}) = 2x_2\mathbf{E}_1 - \frac{1}{6}x_1\mathbf{E}_2 + 3x_3\mathbf{E}_3$$

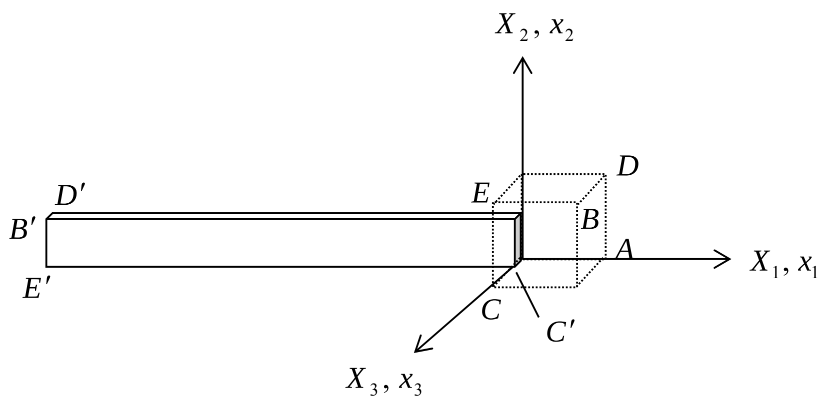


Figure 2.2.3: a deforming cube

Then

$$\mathbf{F} = \frac{\partial x_i}{\partial X_J} = \begin{bmatrix} 0 & -6 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

Once **F** is known, the position of any element can be determined. For example, taking a line element $d\mathbf{X} = [da, 0, 0]^{\mathrm{T}}$, $d\mathbf{x} = \mathbf{F}d\mathbf{X} = [0, da/2, 0]^{\mathrm{T}}$.

■

Homogeneous Deformations

A **homogeneous deformation** is one where the deformation gradient is uniform, i.e. independent of the coordinates, and the associated motion is termed **affine**. Every part of the material deforms as the whole does, and straight parallel lines in the reference configuration map to straight parallel lines in the current configuration, as in the above example. Most examples to be considered in what follows will be of homogeneous deformations; this keeps the algebra to a minimum, but homogeneous deformation analysis is very useful in itself since most of the basic experimental testing of materials, e.g. the uniaxial tensile test, involve homogeneous deformations.

Rigid Body Rotations and Translations

One can add a constant vector **c** to the motion, $\mathbf{x} = \mathbf{x} + \mathbf{c}$, without changing the deformation, $\mathrm{Grad}(\mathbf{x} + \mathbf{c}) = \mathrm{Grad}\mathbf{x}$. Thus the deformation gradient does not take into account rigid-body **translations** of bodies in space. If a body only translates as a rigid body in space, then $\mathbf{F} = \mathbf{I}$, and $\mathbf{x} = \mathbf{X} + \mathbf{c}$ (again, note that **F** does not tell us where in space a particle is, only how it has deformed locally). If there is *no* motion, then not only is $\mathbf{F} = \mathbf{I}$, but $\mathbf{x} = \mathbf{X}$.

If the body rotates as a rigid body (with no translation), then $\mathbf{F} = \mathbf{R}$, a rotation tensor (§1.10.8). For example, for a rotation of θ about the X_2 axis,

$$\mathbf{F} = \begin{bmatrix} \sin\theta & 0 & \cos\theta \\ 0 & 1 & 0 \\ \cos\theta & 0 & -\sin\theta \end{bmatrix}$$

Note that different particles of the same material body can be translating only, rotating only, deforming only, or any combination of these.

The Inverse of the Deformation Gradient

The inverse deformation gradient $\mathbf{F}^{-1}$ carries the spatial line element d**x** to the material line element d**X**. It is defined as

$$\boxed{\mathbf{F}^{-1} = \frac{\partial \mathbf{X}}{\partial \mathbf{x}} = \operatorname{grad}\mathbf{X}, \qquad F_{Ij}^{-1} = \frac{\partial X_I}{\partial x_j}} \quad \textbf{Inverse Deformation Gradient} \qquad (2.2.4)$$

so that

$$\boxed{d\mathbf{X} = \mathbf{F}^{-1}\, d\mathbf{x}, \qquad dX_I = F_{Ij}^{-1}\, dx_j} \quad \textbf{action of } \mathbf{F}^{-1} \qquad (2.2.5)$$

with (see Eqn. 1.15.2)

$$\mathbf{F}^{-1}\mathbf{F} = \mathbf{F}\mathbf{F}^{-1} = \mathbf{I} \qquad F_{iM} F_{Mj}^{-1} = \delta_{ij} \qquad (2.2.6)$$

Cartesian Base Vectors

Explicitly, in terms of the material and spatial base vectors (see 1.14.3),

$$\begin{aligned} \mathbf{F} &= \frac{\partial \mathbf{x}}{\partial X_J} \otimes \mathbf{E}_J = \frac{\partial x_i}{\partial X_J} \mathbf{e}_i \otimes \mathbf{E}_J \\ \mathbf{F}^{-1} &= \frac{\partial \mathbf{X}}{\partial x_j} \otimes \mathbf{e}_j = \frac{\partial X_I}{\partial x_j} \mathbf{E}_I \otimes \mathbf{e}_j \end{aligned} \qquad (2.2.7)$$

so that, for example, $\mathbf{F}d\mathbf{X} = (\partial x_i / \partial X_J)\mathbf{e}_i \otimes \mathbf{E}_J (dX_M \mathbf{E}_M) = (\partial x_i / \partial X_J) dX_J \mathbf{e}_i = d\mathbf{x}$.

Because **F** and $\mathbf{F}^{-1}$ act on vectors in one configuration to produce vectors in the other configuration, they are termed **two-point tensors**. They are defined in both configurations. This is highlighted by their having both reference and current base vectors **E** and **e** in their Cartesian representation 2.2.7.

Here follow some important relations which relate scalar-, vector- and second-order tensor-valued functions in the material and spatial descriptions, the first two relating the material and spatial gradients {▲Problem 1}.

$$\begin{aligned}\mathrm{grad}\phi &= \mathrm{Grad}\phi\,\mathbf{F}^{-1} \\ \mathrm{grad}\mathbf{v} &= \mathrm{Grad}\mathbf{V}\,\mathbf{F}^{-1} \\ \mathrm{div}\mathbf{a} &= \mathrm{Grad}\mathbf{A} : \mathbf{F}^{-\mathrm{T}}\end{aligned} \tag{2.2.8}$$

Here, ϕ is a scalar; $\mathbf{V}$ and $\mathbf{v}$ are the *same* vector, the former being a function of the material coordinates, the material description, the latter a function of the spatial coordinates, the spatial description. Similarly, $\mathbf{A}$ is a second order tensor in the material form and $\mathbf{a}$ is the equivalent spatial form.

The first two of 2.2.8 relate the material gradient to the spatial gradient: the gradient of a function is a measure of how the function changes as one moves through space; since the material coordinates and the spatial coordinates differ, the change in a function with respect to a unit change in the material coordinates will differ from the change in the *same* function with respect to a unit change in the spatial coordinates (see also §2.2.7 below).

Example

Consider the deformation

$$\begin{aligned}\mathbf{x} &= (2X_2 - X_3)\mathbf{e}_1 + (-X_2)\mathbf{e}_2 + (X_1 + 3X_2 + X_3)\mathbf{e}_3 \\ \mathbf{X} &= (x_1 + 5x_2 + x_3)\mathbf{E}_1 + (-x_2)\mathbf{E}_2 + (-x_1 - 2x_2)\mathbf{E}_3\end{aligned}$$

so that

$$\mathbf{F} = \begin{bmatrix} 0 & 2 & -1 \\ 0 & -1 & 0 \\ 1 & 3 & 1 \end{bmatrix}, \qquad \mathbf{F}^{-1} = \begin{bmatrix} 1 & 5 & 1 \\ 0 & -1 & 0 \\ -1 & -2 & 0 \end{bmatrix}$$

Consider the vector $\mathbf{v}(\mathbf{x}) = (2x_1 - x_2)\mathbf{e}_1 + (-3x_2^2 + x_3)\mathbf{e}_2 + (x_1 + x_3)\mathbf{e}_3$ which, in the material description, is

$$\mathbf{V}(\mathbf{X}) = (5X_2 - 2X_3)\mathbf{E}_1 + (X_1 + 3X_2 + X_3 - 3X_2^2)\mathbf{E}_2 + (X_1 + 5X_2)\mathbf{E}_3$$

The material and spatial gradients are

$$\mathrm{Grad}\mathbf{V} = \begin{bmatrix} 0 & 5 & -2 \\ 1 & 3 - 6X_2 & 1 \\ 1 & 5 & 0 \end{bmatrix}, \qquad \mathrm{grad}\mathbf{v} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & -6x_2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

and it can be seen that

$$\mathrm{Grad}\mathbf{V}\,\mathbf{F}^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 6X_2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & -6x_2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \mathrm{grad}\,\mathbf{v}$$

■

2.2.2 The Cauchy-Green Strain Tensors

The deformation gradient describes how a line element in the reference configuration maps into a line element in the current configuration. It has been seen that the deformation gradient gives information about deformation (change of shape) and rigid body rotation, but does not encompass information about possible rigid body translations. The deformation and rigid rotation will be separated shortly (see §2.2.5). To this end, consider the following **strain** tensors; these tensors give direct information about the deformation of the body. Specifically, the **Left Cauchy-Green Strain** and **Right Cauchy-Green Strain** tensors give a measure of how the lengths of line elements and angles between line elements (through the vector dot product) change between configurations.

The Right Cauchy-Green Strain

Consider two line elements in the reference configuration $d\mathbf{X}^{(1)}, d\mathbf{X}^{(2)}$ which are mapped into the line elements $d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}$ in the current configuration. Then, using 1.10.3d,

$$\boxed{\begin{aligned} d\mathbf{x}^{(1)} \cdot d\mathbf{x}^{(2)} &= \left(\mathbf{F}d\mathbf{X}^{(1)}\right)\cdot\left(\mathbf{F}d\mathbf{X}^{(2)}\right) \\ &= d\mathbf{X}^{(1)}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right)d\mathbf{X}^{(2)} \\ &= d\mathbf{X}^{(1)}\mathbf{C}d\mathbf{X}^{(2)} \end{aligned}} \qquad \textbf{action of C} \qquad (2.2.9)$$

where, by definition, $\mathbf{C}$ is the right Cauchy-Green Strain[2]

$$\boxed{\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}, \qquad C_{IJ} = F_{kI}F_{kJ} = \frac{\partial x_k}{\partial X_I}\frac{\partial x_k}{\partial X_J}} \qquad \textbf{Right Cauchy-Green Strain} \qquad (2.2.10)$$

It is a symmetric, positive definite (which will be clear from Eqn. 2.2.17 below), tensor, which implies that it has real positive eigenvalues (*cf.* §1.11.2), and this has important consequences (see later). Explicitly in terms of the base vectors,

$$\mathbf{C} = \left(\frac{\partial x_k}{\partial X_I}\mathbf{E}_I \otimes \mathbf{e}_k\right)\left(\frac{\partial x_m}{\partial X_J}\mathbf{e}_m \otimes \mathbf{E}_J\right) = \frac{\partial x_k}{\partial X_I}\frac{\partial x_k}{\partial X_J}\mathbf{E}_I \otimes \mathbf{E}_J \,. \qquad (2.2.11)$$

Just as the line element $d\mathbf{X}$ is a vector defined in and associated with the reference configuration, $\mathbf{C}$ is defined in and associated with the reference configuration, acting on vectors in the reference configuration, and so is called a **material tensor**.

---

[2] "right" because **F** is on the right of the formula

The inverse of **C**, $\mathbf{C}^{-1}$, is called the **Piola deformation tensor**.

The Left Cauchy-Green Strain

Consider now the following, using Eqn. 1.10.18c:

$$\boxed{\begin{aligned} d\mathbf{X}^{(1)} \cdot d\mathbf{X}^{(2)} &= \left(\mathbf{F}^{-1} d\mathbf{x}^{(1)}\right) \cdot \left(\mathbf{F}^{-1} d\mathbf{x}^{(2)}\right) \\ &= d\mathbf{x}^{(1)} \left(\mathbf{F}^{-\mathrm{T}} \mathbf{F}^{-1}\right) d\mathbf{x}^{(2)} \\ &= d\mathbf{x}^{(1)} \mathbf{b}^{-1} d\mathbf{x}^{(2)} \end{aligned}} \qquad \textbf{action of } \mathbf{b}^{-1} \qquad (2.2.12)$$

where, by definition, **b** is the left Cauchy-Green Strain, also known as the **Finger tensor**:

$$\boxed{\mathbf{b} = \mathbf{F}\mathbf{F}^{\mathrm{T}}, \qquad b_{ij} = F_{iK} F_{jK} = \frac{\partial x_i}{\partial X_K} \frac{\partial x_j}{\partial X_K}} \qquad \textbf{Left Cauchy-Green Strain} \qquad (2.2.13)$$

Again, this is a symmetric, positive definite tensor, only here, **b** is defined in the current configuration and so is called a **spatial tensor**.

The inverse of **b**, $\mathbf{b}^{-1}$, is called the **Cauchy deformation tensor**.

It can be seen that the right and left Cauchy-Green tensors are related through

$$\mathbf{C} = \mathbf{F}^{-1}\mathbf{b}\mathbf{F}, \qquad \mathbf{b} = \mathbf{F}\mathbf{C}\mathbf{F}^{-1} \qquad (2.2.14)$$

Note that tensors can be material (e.g. **C**), two-point (e.g. **F**) or spatial (e.g. **b**). Whatever type they are, they can always be described using material or spatial coordinates through the motion mapping 2.1.3, that is, using the material or spatial descriptions. Thus one distinguishes between, for example, a spatial tensor, which is an intrinsic property of a tensor, and the spatial description of a tensor.

The Principal Scalar Invariants of the Cauchy-Green Tensors

Using 1.10.10b,

$$\mathrm{tr}\mathbf{C} = \mathrm{tr}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F}\right) = \mathrm{tr}\left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right) = \mathrm{tr}\mathbf{b} \qquad (2.2.15)$$

This holds also for arbitrary powers of these tensors, $\mathrm{tr}\mathbf{C}^n = \mathrm{tr}\mathbf{b}^n$, and therefore, from Eqn. 1.11.17, the invariants of **C** and **b** are equal.

2.2.3 The Stretch

The **stretch** (or the **stretch ratio**) λ is defined as the ratio of the length of a deformed line element to the length of the corresponding undeformed line element:

$$\boxed{\lambda = \frac{|d\mathbf{x}|}{|d\mathbf{X}|}} \qquad \textbf{The Stretch} \tag{2.2.16}$$

From the relations involving the Cauchy-Green Strains, letting $d\mathbf{X}^{(1)} = d\mathbf{X}^{(2)} \equiv d\mathbf{X}$, $d\mathbf{x}^{(1)} = d\mathbf{x}^{(2)} \equiv d\mathbf{x}$, and dividing across by the square of the length of $d\mathbf{X}$ or $d\mathbf{x}$,

$$\lambda^2 = \left(\frac{|d\mathbf{x}|}{|d\mathbf{X}|}\right)^2 = d\hat{\mathbf{X}}\mathbf{C}d\hat{\mathbf{X}}, \qquad \lambda^{-2} = \left(\frac{|d\mathbf{X}|}{|d\mathbf{x}|}\right)^2 = d\hat{\mathbf{x}}\mathbf{b}^{-1}d\hat{\mathbf{x}} \tag{2.2.17}$$

Here, the quantities $d\hat{\mathbf{X}} = d\mathbf{X}/|d\mathbf{X}|$ and $d\hat{\mathbf{x}} = d\mathbf{x}/|d\mathbf{x}|$ are unit vectors in the directions of $d\mathbf{X}$ and $d\mathbf{x}$. Thus, through these relations, **C** and **b** determine how much a line element stretches (and, from 2.2.17, **C** and **b** can be seen to be indeed positive definite).

One says that a line element is **extended**, **unstretched** or **compressed** according to $\lambda > 1$, $\lambda = 1$ or $\lambda < 1$.

Stretching along the Coordinate Axes

Consider three line elements lying along the three coordinate axes[3]. Suppose that the material deforms in a special way, such that these line elements undergo a **pure stretch**, that is, they change length with no change in the right angles between them. If the stretches in these directions are λ_1, λ_2 and λ_3, then

$$x_1 = \lambda_1 X_1, \qquad x_2 = \lambda_2 X_2, \qquad x_3 = \lambda_3 X_3 \tag{2.2.18}$$

and the deformation gradient has only diagonal elements in its matrix form:

$$\mathbf{F} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \qquad F_{iJ} = \lambda_i \delta_{iJ} \quad \text{(no sum)} \tag{2.2.19}$$

Whereas material undergoes pure stretch along the coordinate directions, line elements off-axes will in general stretch/contract *and* rotate relative to each other. For example, a line element $d\mathbf{X} = [\alpha, \alpha, 0]^{\mathrm{T}}$ stretches by $\lambda = \sqrt{d\hat{\mathbf{X}}\mathbf{C}d\hat{\mathbf{X}}} = \sqrt{\left(\lambda_1^2 + \lambda_2^2\right)/2}$ with $d\mathbf{x} = [\lambda_1\alpha, \lambda_2\alpha, 0]^{\mathrm{T}}$, and rotates if $\lambda_1 \neq \lambda_2$.

It will be shown below that, for any deformation, there are always three mutually orthogonal directions along which material undergoes a pure stretch. These directions, the coordinate axes in this example, are called the **principal axes** of the material and the associated stretches are called the **principal stretches**.

[3] with the material and spatial basis vectors coincident

The Case of F Real and Symmetric

Consider now another special deformation, where **F** is a real symmetric tensor, in which case the eigenvalues are real and the eigenvectors form an orthonormal basis (*cf.* §1.11.2)[4]. In any given coordinate system, **F** will in general result in the stretching of line elements and the changing of the angles between line elements. However, if one chooses a coordinate set to be the eigenvectors of **F**, then from Eqn. 1.11.11-12 one can write[5]

$$\mathbf{F} = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{N}}_i, \qquad [\mathbf{F}] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \tag{2.2.20}$$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of **F**. The eigenvalues are the principal stretches and the eigenvectors are the principal axes. This indicates that as long as **F** is real and symmetric, one can always find a coordinate system along whose axes the material undergoes a pure stretch, with no rotation. This topic will be discussed more fully in §2.2.5 below.

2.2.4 The Green-Lagrange and Euler-Almansi Strain Tensors

Whereas the left and right Cauchy-Green tensors give information about the change in angle between line elements and the stretch of line elements, the **Green-Lagrange strain** and the **Euler-Almansi strain** tensors directly give information about the change in the squared length of elements.

Specifically, when the Green-Lagrange strain **E** operates on a line element d**X**, it gives (half) the change in the squares of the undeformed and deformed lengths:

$$\boxed{\begin{aligned} \frac{|d\mathbf{x}|^2 - |d\mathbf{X}|^2}{2} &= \frac{1}{2}\{d\mathbf{X}\mathbf{C}d\mathbf{X} - d\mathbf{X}\cdot d\mathbf{X}\} \\ &= \frac{1}{2}\{d\mathbf{X}(\mathbf{C}-\mathbf{I})d\mathbf{X}\} \\ &\equiv d\mathbf{X}\mathbf{E}d\mathbf{X} \end{aligned}} \qquad \textbf{action of E} \tag{2.2.21}$$

where

$$\boxed{\mathbf{E} = \frac{1}{2}(\mathbf{C}-\mathbf{I}) = \frac{1}{2}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F}-\mathbf{I}\right), \quad E_{IJ} = \frac{1}{2}\left(C_{IJ} - \delta_{IJ}\right)} \qquad \textbf{Green-Lagrange Strain} \tag{2.2.22}$$

It is a symmetric positive definite material tensor. Similarly, the (symmetric spatial) Euler-Almansi strain tensor is defined through

[4] in fact, **F** in this case will have to be positive definite, with $\det \mathbf{F} > 0$ (see later in §2.2.8)

[5] $\hat{\mathbf{n}}_i$ are the eigenvectors for the basis $\mathbf{e}_i$, $\hat{\mathbf{N}}_I$ for the basis $\hat{\mathbf{E}}_i$, with $\hat{\mathbf{n}}_i$, $\hat{\mathbf{N}}_I$ coincident; when the bases are not coincident, the notion of rotating line elements becomes ambiguous – this topic will be examined later in the context of *objectivity*

$$\boxed{\frac{|d\mathbf{x}|^2 - |d\mathbf{X}|^2}{2} = d\mathbf{x}\,\mathbf{e}\,d\mathbf{x}} \qquad \textbf{action of e} \tag{2.2.23}$$

and

$$\boxed{\mathbf{e} = \frac{1}{2}\left(\mathbf{I} - \mathbf{b}^{-1}\right) = \frac{1}{2}\left(\mathbf{I} - \mathbf{F}^{-\mathrm{T}}\mathbf{F}^{-1}\right)} \qquad \textbf{Euler-Almansi Strain} \tag{2.2.24}$$

Physical Meaning of the Components of E

Take a line element in the 1-direction, $d\mathbf{X}_{(1)} = [dX_1,\ 0,\ 0]^{\mathrm{T}}$, so that $d\hat{\mathbf{X}}_{(1)} = [1,\ 0,\ 0]^{\mathrm{T}}$. The square of the stretch of this element is

$$\lambda_{(1)}^2 = d\hat{\mathbf{X}}_{(1)}\mathbf{C}d\hat{\mathbf{X}}_{(1)} = C_{11} \quad \to \quad E_{11} = \frac{1}{2}\left(C_{11} - 1\right) = \frac{1}{2}\left(\lambda_{(1)}^2 - 1\right)$$

The unit extension is $\left(|d\mathbf{x}| - |d\mathbf{X}|\right)/|d\mathbf{X}| = \lambda - 1$. Denoting the unit extension of $d\mathbf{X}_{(1)}$ by $\mathbf{E}_{(1)}$, one has

$$E_{11} = \mathbf{E}_{(1)} + \frac{1}{2}\mathbf{E}_{(1)}^2 \tag{2.2.25}$$

and similarly for the other diagonal elements E_{22}, E_{33}.

When the deformation is small, $\mathbf{E}_{(1)}^2$ is small in comparison to $\mathbf{E}_{(1)}$, so that $E_{11} \approx \mathbf{E}_{(1)}$. For small deformations then, the diagonal terms are equivalent to the unit extensions.

Let θ_{12} denote the angle between the deformed elements which were initially parallel to the X_1 and X_2 axes. Then

$$\begin{aligned}\cos\theta_{12} &= \frac{d\mathbf{x}_{(1)}}{|d\mathbf{x}_{(1)}|} \cdot \frac{d\mathbf{x}_{(2)}}{|d\mathbf{x}_{(2)}|} = \frac{|d\mathbf{X}_{(1)}||d\mathbf{X}_{(2)}|}{|d\mathbf{x}_{(1)}||d\mathbf{x}_{(2)}|}\left\{\frac{d\mathbf{X}_{(1)}}{|d\mathbf{X}_{(1)}|} \cdot \mathbf{C}\frac{d\mathbf{X}_{(2)}}{|d\mathbf{X}_{(2)}|}\right\} = \frac{C_{12}}{\lambda_{(1)}\lambda_{(2)}} \\ &= \frac{2E_{12}}{\sqrt{2E_{11} + 1}\sqrt{2E_{22} + 1}}\end{aligned} \tag{2.2.26}$$

and similarly for the other off-diagonal elements. Note that if $\theta_{12} = \pi/2$, so that there is no angle change, then $E_{12} = 0$. Again, if the deformation is small, then E_{11}, E_{22} are small, and

$$\frac{\pi}{2} - \theta_{12} \approx \sin\left(\frac{\pi}{2} - \theta_{12}\right) = \cos\theta_{12} \approx 2E_{12} \tag{2.2.27}$$

In words: for small deformations, the component E_{12} gives half the change in the original right angle.

2.2.5 Stretch and Rotation Tensors

The deformation gradient can always be decomposed into the product of two tensors, a stretch tensor and a rotation tensor (in one of two different ways, material or spatial versions). This is known as the **polar decomposition**, and is discussed in §1.11.7. One has

$$\boxed{\mathbf{F} = \mathbf{R}\mathbf{U}} \qquad \textbf{Polar Decomposition (Material)} \tag{2.2.28}$$

Here, **R** is a proper orthogonal tensor, i.e. $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$ with $\det \mathbf{R} = 1$, called *the* **rotation tensor**. It is a measure of the local rotation at **X**.

The decomposition is not unique; it is made unique by choosing **U** to be a *symmetric* tensor, called the **right stretch tensor**. It is a measure of the local stretching (or contraction) of material at **X**. Consider a line element $d\mathbf{X}$. Then

$$\lambda d\hat{\mathbf{x}} = \mathbf{F}d\hat{\mathbf{X}} = \mathbf{R}\mathbf{U}d\hat{\mathbf{X}} \tag{2.2.29}$$

and so {▲Problem 2}

$$\lambda^2 = d\hat{\mathbf{X}}\mathbf{U} \cdot \mathbf{U}d\hat{\mathbf{X}} \tag{2.2.30}$$

Thus (this is a definition of **U**)

$$\boxed{\mathbf{U} = \sqrt{\mathbf{C}} \qquad (\mathbf{C} = \mathbf{U}\mathbf{U})} \qquad \textbf{The Right Stretch Tensor} \tag{2.2.31}$$

From 2.2.30, the right Cauchy-Green strain **C** (and by consequence the Euler-Lagrange strain **E**) only give information about the stretch of line elements; it does not give information about the rotation that is experienced by a particle during motion. The deformation gradient **F**, however, contains information about both the stretch and rotation. It can also be seen from 2.2.30-1 that **U** is a material tensor.

Note that, since

$$d\mathbf{x} = \mathbf{R}(\mathbf{U}d\mathbf{X}),$$

the undeformed line element is *first* stretched by **U** and is *then* rotated by **R** into the deformed element $d\mathbf{x}$ (the element may also undergo a rigid body translation **c**), Fig. 2.2.4. **R** is a two-point tensor.

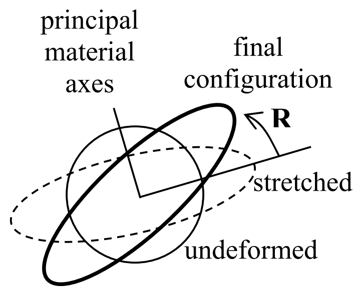


Figure 2.2.4: the polar decomposition

Evaluation of U

In order to evaluate **U**, it is necessary to evaluate $\sqrt{\mathbf{C}}$. To evaluate the square-root, **C** must first be obtained in relation to its principal axes, so that it is diagonal, and then the square root can be taken of the diagonal elements, since its eigenvalues will be positive (see §1.11.6). Then the tensor needs to be transformed back to the original coordinate system.

Example

Consider the motion

$$x_1 = 2X_1 - 2X_2, \quad x_2 = X_1 + X_2, \quad x_3 = X_3$$

The (homogeneous) deformation of a unit square in the $x_1 - x_2$ plane is as shown in Fig. 2.2.5.

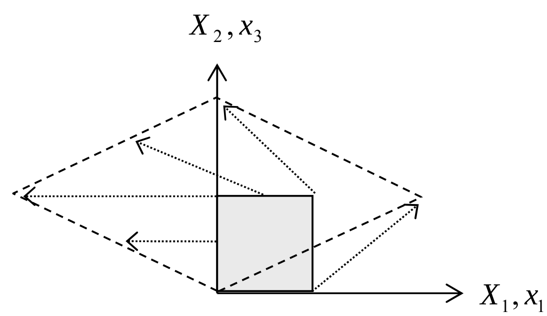


Figure 2.2.5: deformation of a square

One has

$$[\mathbf{F}] = \begin{bmatrix} 2 & -2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{basis}: (\mathbf{e}_i \otimes \mathbf{E}_j), \qquad [\mathbf{C}] = [\mathbf{F}^{\mathrm{T}}\mathbf{F}] = \begin{bmatrix} 5 & -3 & 0 \\ -3 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{basis}: (\mathbf{E}_i \otimes \mathbf{E}_j)$$

Note that **F** is not symmetric, so that it might have only one real eigenvalue (in fact here it does have complex eigenvalues), and the eigenvectors may not be orthonormal. **C**, on the other hand, by its very definition, *is* symmetric; it is in fact positive definite and so has positive real eigenvalues forming an orthonormal set.

To determine the principal axes of **C**, it is necessary to evaluate the eigenvalues/eigenvectors of the tensor. The eigenvalues are the roots of the characteristic equation 1.11.5,

$$\alpha^3 - \mathrm{I}_{\mathbf{C}}\alpha^2 + \mathrm{II}_{\mathbf{C}}\alpha - \mathrm{III}_{\mathbf{C}} = 0$$

and the first, second and third invariants of the tensor are given by 1.11.6 so that $\alpha^3 - 11\alpha^2 + 26\alpha - 16 = 0$, with roots $\alpha = 8, 2, 1$. The three corresponding eigenvectors are found from 1.11.8,

$$\begin{aligned} (C_{11}-\alpha)\hat{N}_1 + C_{12}\hat{N}_2 + C_{13}\hat{N}_3 &= 0 \\ C_{21}\hat{N}_1 + (C_{22}-\alpha)\hat{N}_2 + C_{23}\hat{N}_3 &= 0 \\ C_{31}\hat{N}_1 + C_{32}\hat{N}_2 + (C_{33}-\alpha)\hat{N}_3 &= 0 \end{aligned} \quad \rightarrow \quad \begin{aligned} (5-\alpha)\hat{N}_1 - 3\hat{N}_2 &= 0 \\ -3\hat{N}_1 + (5-\alpha)\hat{N}_2 &= 0 \\ (1-\alpha)\hat{N}_3 &= 0 \end{aligned}$$

Thus (normalizing the eigenvectors so that they are unit vectors, and form a right-handed set, Fig. 2.2.6):

(i) for $\alpha = 8$, $-3\hat{N}_1 - 3\hat{N}_2 = 0, -3\hat{N}_1 - 3\hat{N}_2 = 0, -7\hat{N}_3 = 0$, $\quad \hat{\mathbf{N}}_1 = \frac{1}{\sqrt{2}}\mathbf{E}_1 - \frac{1}{\sqrt{2}}\mathbf{E}_2$

(ii) for $\alpha = 2$, $3\hat{N}_1 - 3\hat{N}_2 = 0, -3\hat{N}_1 + 3\hat{N}_2 = 0, -\hat{N}_3 = 0$, $\quad \hat{\mathbf{N}}_2 = \frac{1}{\sqrt{2}}\mathbf{E}_1 + \frac{1}{\sqrt{2}}\mathbf{E}_2$

(iii) for $\alpha = 1$, $4\hat{N}_1 - 3\hat{N}_2 = 0, -3\hat{N}_1 + 4\hat{N}_2 = 0, 0\hat{N}_3 = 0$, $\quad \hat{\mathbf{N}}_3 = \mathbf{E}_3$

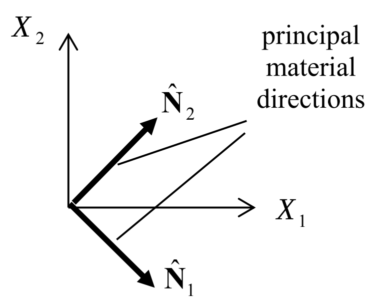


Figure 2.2.6: deformation of a square

Thus the right Cauchy-Green strain tensor **C**, with respect to coordinates with base vectors $\mathbf{E}_1' = \hat{\mathbf{N}}_1$, $\mathbf{E}_2' = \hat{\mathbf{N}}_2$ and $\mathbf{E}_3' = \hat{\mathbf{N}}_3$, that is, in terms of principal coordinates, is

$$[\mathbf{C}] = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{basis}: \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_j$$

This result can be checked using the tensor transformation formulae 1.13.6, $[\mathbf{C}'] = [\mathbf{Q}]^{\mathrm{T}}[\mathbf{C}][\mathbf{Q}]$, where Q is the transformation matrix of direction cosines (see also the example at the end of §1.5.2),

$$Q_{ij} = \begin{bmatrix} \mathbf{e}_1 \cdot \mathbf{e}'_1 & \mathbf{e}_1 \cdot \mathbf{e}'_2 & \mathbf{e}_1 \cdot \mathbf{e}'_3 \\ \mathbf{e}_2 \cdot \mathbf{e}'_1 & \mathbf{e}_2 \cdot \mathbf{e}'_2 & \mathbf{e}_2 \cdot \mathbf{e}'_3 \\ \mathbf{e}_3 \cdot \mathbf{e}'_1 & \mathbf{e}_3 \cdot \mathbf{e}'_2 & \mathbf{e}_3 \cdot \mathbf{e}'_3 \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \hat{\mathbf{N}}_1 & \hat{\mathbf{N}}_2 & \hat{\mathbf{N}}_3 \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The stretch tensor **U**, with respect to the principal directions is

$$[\mathbf{U}] = \left[\sqrt{\mathbf{C}}\right] = \begin{bmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad \text{basis} : \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_j$$

These eigenvalues of **U** (which are the square root of those of **C**) are the principal stretches and, as before, they are labeled $\lambda_1, \lambda_2, \lambda_3$.

In the original coordinate system, using the inverse tensor transformation rule 1.13.6, $[\mathbf{U}] = [\mathbf{Q}][\mathbf{U}'][\mathbf{Q}]^{\mathrm{T}}$,

$$[\mathbf{U}] = \begin{bmatrix} 3/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 3/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{basis} : \mathbf{E}_i \otimes \mathbf{E}_j$$

so that

$$[\mathbf{R}] = \left[\mathbf{F}\mathbf{U}^{-1}\right] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{basis} : \mathbf{e}_i \otimes \mathbf{E}_j$$

and it can be verified that **R** is a rotation tensor, i.e. is proper orthogonal.

Returning to the deformation of the unit square, the stretch and rotation are as illustrated in Fig. 2.2.7 – the action of **U** is indicated by the arrows, deforming the unit square to the dotted parallelogram, whereas **R** rotates the parallelogram through 45° as a rigid body to its final position.

Note that the line elements along the diagonals (indicated by the heavy lines) lie along the principal directions of **U** and therefore undergo a pure stretch; the diagonal in the $\hat{\mathbf{N}}_1$ direction has stretched but has also moved with a rigid translation.

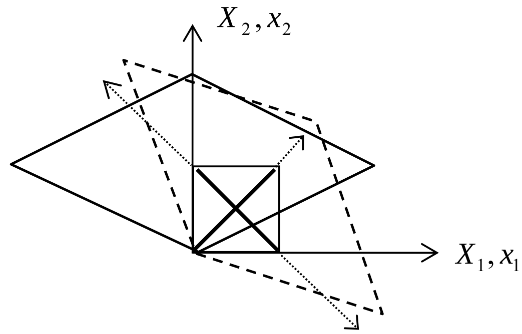


Figure 2.2.7: stretch and rotation of a square

■

Spatial Description

A polar decomposition can be made in the spatial description. In that case,

$$\boxed{\mathbf{F} = \mathbf{v}\mathbf{R}} \qquad \textbf{Polar Decomposition (Spatial)} \tag{2.2.32}$$

Here **v** is a symmetric, positive definite second order tensor called the **left stretch tensor**, and $\mathbf{vv} = \mathbf{b}$, where **b** is the left Cauchy-Green tensor. **R** is the same rotation tensor as appears in the material description. Thus an elemental sphere can be regarded as first stretching into an ellipsoid, whose axes are the principal material axes (the principal axes of **U**), and then rotating; or first rotating, and then stretching into an ellipsoid whose axes are the **principal spatial axes** (the principal axes of **v**). The end result is the same.

The development in the spatial description is similar to that given above for the material description, and one finds by analogy with 2.2.30,

$$\lambda^{-2} = d\hat{\mathbf{x}}\mathbf{v}^{-1} \cdot \mathbf{v}^{-1} d\hat{\mathbf{x}} \tag{2.2.33}$$

In the above example, it turns out that **v** takes the simple diagonal form

$$[\mathbf{v}] = \begin{bmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{basis}: \mathbf{e}_i \otimes \mathbf{e}_j \,.$$

so the unit square rotates first and then undergoes a pure stretch along the coordinate axes, which are the principal spatial axes, and the sequence is now as shown in Fig. 2.2.9.

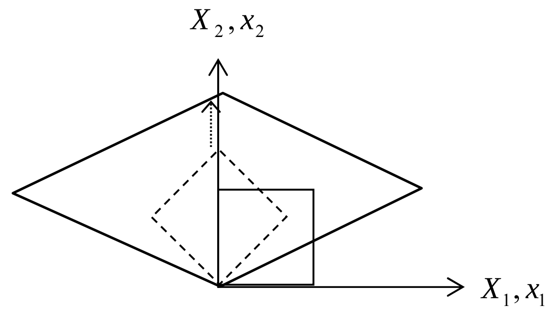


Figure 2.2.8: stretch and rotation of a square in spatial description

Relationship between the Material and Spatial Decompositions

Comparing the two decompositions, one sees that the material and spatial tensors involved are related through

$$\mathbf{v} = \mathbf{R}\mathbf{U}\mathbf{R}^{\mathrm{T}}, \quad \mathbf{b} = \mathbf{R}\mathbf{C}\mathbf{R}^{\mathrm{T}} \tag{2.2.34}$$

Further, suppose that **U** has an eigenvalue λ and an eigenvector $\hat{\mathbf{N}}$. Then $\mathbf{U}\hat{\mathbf{N}} = \lambda\hat{\mathbf{N}}$, so that $\mathbf{RUN} = \lambda\mathbf{RN}$. But $\mathbf{RU} = \mathbf{vR}$, so $\mathbf{v}\left(\mathbf{R}\hat{\mathbf{N}}\right) = \lambda\left(\mathbf{R}\hat{\mathbf{N}}\right)$. Thus **v** also has an eigenvalue λ, but an eigenvector $\hat{\mathbf{n}} = \mathbf{R}\hat{\mathbf{N}}$. From this, it is seen that the rotation tensor **R** maps the principal material axes into the principal spatial axes. It also follows that **R** and **F** can be written explicitly in terms of the material and spatial principal axes (compare the first of these with 1.10.25)[6]:

$$\mathbf{R} = \hat{\mathbf{n}}_i \otimes \hat{\mathbf{N}}_i, \qquad \mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{R}\sum_{i=1}^{3} \lambda_i \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{N}}_i \tag{2.2.35}$$

and the deformation gradient acts on the principal axes base vectors according to {▲Problem 4}

$$\mathbf{F}\hat{\mathbf{N}}_i = \lambda_i \hat{\mathbf{n}}_i, \qquad \mathbf{F}^{-\mathrm{T}}\hat{\mathbf{N}}_i = \frac{1}{\lambda_i}\hat{\mathbf{n}}_i, \qquad \mathbf{F}^{-1}\hat{\mathbf{n}}_i = \frac{1}{\lambda_i}\hat{\mathbf{N}}_i, \qquad \mathbf{F}^{\mathrm{T}}\hat{\mathbf{n}}_i = \lambda_i \hat{\mathbf{N}}_i \tag{2.2.36}$$

The representation of **F** and **R** in terms of both material and spatial principal base vectors in 2.3.35 highlights their two-point character.

Other Strain Measures

Some other useful measures of strain are

The **Hencky strain** measure: $\mathbf{H} \equiv \ln \mathbf{U}$ (material) or $\mathbf{h} = \ln \mathbf{v}$ (spatial)

[6] this is not a spectral decomposition of **F** (unless **F** happens to be symmetric, which it must be in order to have a spectral decomposition)

The **Biot strain** measure: $\overline{\mathbf{B}} = \mathbf{U} - \mathbf{I}$ (material) or $\overline{\mathbf{b}} = \mathbf{v} - \mathbf{I}$ (spatial)

The Hencky strain is evaluated by first evaluating **U** along the principal axes, so that the logarithm can be taken of the diagonal elements.

The material tensors **H**, $\overline{\mathbf{B}}$, **C**, **U** and **E** are coaxial tensors, with the same eigenvectors $\hat{\mathbf{N}}_i$. Similarly, the spatial tensors **h**, $\overline{\mathbf{b}}$, **b**, **v** and **e** are coaxial with the same eigenvectors $\hat{\mathbf{n}}_i$. From the definitions, the spectral decompositions of these tensors are

$$\begin{aligned}
\mathbf{U} &= \sum_{i=1}^{3} \lambda_i \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i & \mathbf{v} &= \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \\
\mathbf{C} &= \sum_{i=1}^{3} \lambda_i^2 \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i & \mathbf{b} &= \sum_{i=1}^{3} \lambda_i^2 \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \\
\mathbf{E} &= \sum_{i=1}^{3} \tfrac{1}{2}\left(\lambda_i^2 - 1\right) \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i & \mathbf{e} &= \sum_{i=1}^{3} \tfrac{1}{2}\left(1 - 1/\lambda_i^2\right) \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \\
\mathbf{H} &= \sum_{i=1}^{3} \left(\ln \lambda_i\right) \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i & \mathbf{h} &= \sum_{i=1}^{3} \left(\ln \lambda_i\right) \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \\
\overline{\mathbf{B}} &= \sum_{i=1}^{3} \left(\lambda_i - 1\right) \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i & \overline{\mathbf{b}} &= \sum_{i=1}^{3} \left(\lambda_i - 1\right) \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i
\end{aligned} \tag{2.2.37}$$

Deformation of a Circular Material Element

A circular material element will deform into an ellipse, as indicated in Figs. 2.2.2 and 2.2.4. This can be shown as follows. With respect to the principal axes, an undeformed line element $d\mathbf{X} = dX_1\mathbf{N}_1 + dX_2\mathbf{N}_2$ has magnitude squared $\left(dX_1\right)^2 + \left(dX_2\right)^2 = c^2$, where c is the radius of the circle, Fig. 2.2.9. The deformed element is $d\mathbf{x} = \mathbf{U}d\mathbf{X}$, or $d\mathbf{x} = \lambda_1 dX_1\mathbf{N}_1 + \lambda_2 dX_2\mathbf{N}_2 \equiv dx_1\mathbf{n}_1 + dx_2\mathbf{n}_2$. Thus $dx_1/\lambda_1 = dX_1$, $dx_2/\lambda_2 = dX_2$, which leads to the standard equation of an ellipse with major and minor axes $\lambda_1 c, \lambda_2 c$: $\left(dx_1/\lambda_1 c\right)^2 + \left(dx_2/\lambda_2 c\right)^2 = 1$.

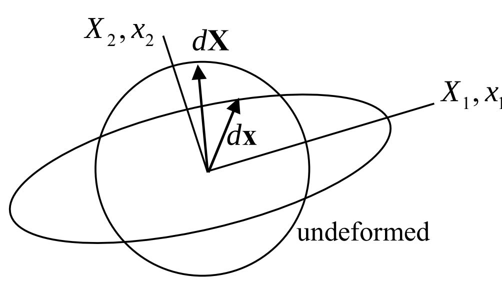


Figure 2.2.9: a circular element deforming into an ellipse

2.2.6 Some Simple Deformations

In this section, some elementary deformations are considered.

Pure Stretch

This deformation has already been seen, but now it can be viewed as a special case of the polar decomposition. The motion is

$$\boxed{x_1 = \lambda_1 X_1, \quad x_2 = \lambda_2 X_2, \quad x_3 = \lambda_3 X_3} \qquad \textbf{Pure Stretch} \qquad (2.2.38)$$

and the deformation gradient is

$$\mathbf{F} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \mathbf{RU} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Here, $\mathbf{R} = \mathbf{I}$ and there is no rotation. $\mathbf{U} = \mathbf{F}$ and the principal material axes are coincident with the material coordinate axes. $\lambda_1, \lambda_2, \lambda_3$, the eigenvalues of **U**, are the principal stretches.

Stretch with rotation

Consider the motion

$$x_1 = X_1 - kX_2, \quad x_2 = kX_1 + X_2, \quad x_3 = X_3$$

so that

$$\mathbf{F} = \begin{bmatrix} 1 & -k & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{RU} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sec\theta & 0 & 0 \\ 0 & \sec\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where $k = \tan\theta$. This decomposition shows that the deformation consists of material stretching by $\sec\theta (= \sqrt{1+k^2})$, the principal stretches, along each of the axes, followed by a rigid body rotation through an angle θ about the $X_3 = 0$ axis, Fig. 2.2.10. The deformation is relatively simple because the principal material axes are aligned with the material coordinate axes (so that **U** is diagonal). The deformation of the unit square is as shown in Fig. 2.2.10.

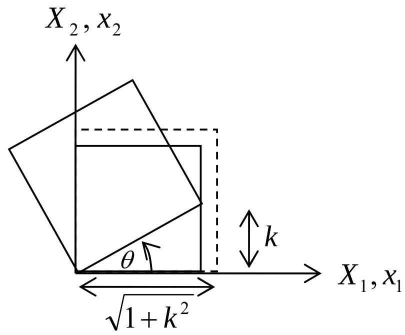


Figure 2.2.10: stretch with rotation

Pure Shear

Consider the motion

$$\boxed{x_1 = X_1 + kX_2, \quad x_2 = kX_1 + X_2, \quad x_3 = X_3} \qquad \textbf{Pure Shear} \qquad (2.2.39)$$

so that

$$\mathbf{F} = \begin{bmatrix} 1 & k & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{RU} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & k & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where, since **F** is symmetric, there is no rotation, and $\mathbf{F} = \mathbf{U}$. Since the rotation is zero, one can work directly with **U** and not have to consider **C**. The eigenvalues of **U**, the principal stretches, are $1+k, 1-k, 1$, with corresponding principal directions $\hat{\mathbf{N}}_1 = \frac{1}{\sqrt{2}}\mathbf{E}_1 + \frac{1}{\sqrt{2}}\mathbf{E}_2$, $\hat{\mathbf{N}}_2 = -\frac{1}{\sqrt{2}}\mathbf{E}_1 + \frac{1}{\sqrt{2}}\mathbf{E}_2$ and $\hat{\mathbf{N}}_3 = \mathbf{E}_3$.

The deformation of the unit square is as shown in Fig. 2.2.11. The diagonal indicated by the heavy line stretches by an amount $1+k$ whereas the other diagonal contracts by an amount $1-k$. An element of material along the diagonal will undergo a pure stretch as indicated by the stretching of the dotted box.

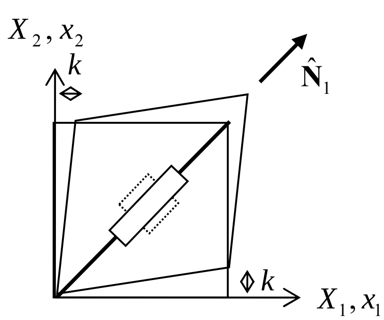


Figure 2.2.11: pure shear

Simple Shear

Consider the motion

$$\boxed{x_1 = X_1 + kX_2, \quad x_2 = X_2, \quad x_3 = X_3} \quad \textbf{Simple Shear} \tag{2.2.40}$$

so that

$$\mathbf{F} = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & k & 0 \\ k & 1+k^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The invariants of **C** are $\mathrm{I}_\mathbf{C} = 3 + k^2$, $\mathrm{II}_\mathbf{C} = 3 + k^2$, $\mathrm{III}_\mathbf{C} = 1$ and the characteristic equation is $\lambda^3 + (3 + k^2)\lambda(1-\lambda) - 1 = 0$, so the principal values of **C** are $\lambda = 1 + \frac{1}{2}k^2 \pm \frac{1}{2}k\sqrt{4 + k^2}, 1$. The principal values of **U** are the (positive) square-roots of these: $\lambda = \frac{1}{2}\sqrt{4 + k^2} \pm \frac{1}{2}k, 1$. These can be written as $\lambda = \sec\theta \pm \tan\theta, 1$ by letting $\tan\theta = \frac{1}{2}k$. The corresponding eigenvectors of **C** are

$$\hat{\mathbf{N}}_1 = \frac{k}{\frac{1}{2}k^2 + \frac{1}{2}k\sqrt{4+k^2}}\mathbf{E}_1 + \mathbf{E}_2, \quad \hat{\mathbf{N}}_2 = \frac{k}{\frac{1}{2}k^2 - \frac{1}{2}k\sqrt{4+k^2}}\mathbf{E}_1 + \mathbf{E}_2, \quad \hat{\mathbf{N}}_3 = \mathbf{E}_3$$

or, normalizing so that they are of unit size, and writing in terms of θ,

$$\hat{\mathbf{N}}_1 = \sqrt{\frac{1-\sin\theta}{2}}\mathbf{E}_1 + \sqrt{\frac{1+\sin\theta}{2}}\mathbf{E}_2, \quad \hat{\mathbf{N}}_2 = -\sqrt{\frac{1+\sin\theta}{2}}\mathbf{E}_1 + \sqrt{\frac{1-\sin\theta}{2}}\mathbf{E}_2, \quad \hat{\mathbf{N}}_3 = \mathbf{E}_3$$

The transformation matrix of direction cosines is then

$$[\mathbf{Q}] = \begin{bmatrix} \sqrt{(1-\sin\theta)/2} & -\sqrt{(1+\sin\theta)/2} & 0 \\ \sqrt{(1+\sin\theta)/2} & \sqrt{(1-\sin\theta)/2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so that, using the inverse transformation formula, $[\mathbf{U}] = [\mathbf{Q}][\mathbf{U}'][\mathbf{Q}]^{\mathrm{T}}$, one obtains **U** in terms of the original coordinates, and hence

$$\mathbf{F} = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{RU} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & (1+\sin^2\theta)/\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The deformation of the unit square is shown in Fig. 2.2.12 (for $k = 0.2, \theta = 5.71^{\mathrm{o}}$). The square first undergoes a pure stretch/contraction ($\hat{\mathbf{N}}_1$ is in this case at 47.86^{o} to the X_1

axis, with the diagonal of the square becoming the diagonal of the parallelogram, at 45.5° to the X_1 axis), and is then brought to its final position by a negative (clockwise) rotation of θ.

For this deformation, $\det \mathbf{F} = 1$ and, as will be shown below, this means that the simple shear deformation is volume-preserving.

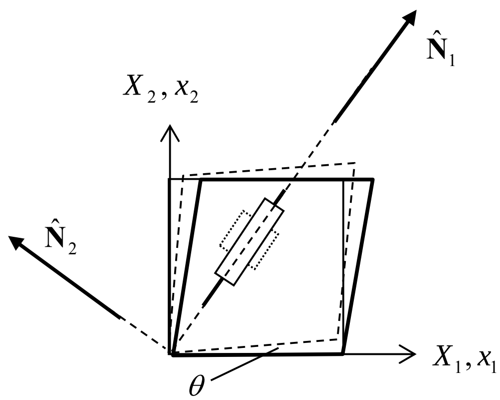


Figure 2.2.12: simple shear

2.2.7 Displacement & Displacement Gradients

The displacement of a material particle[7] is the movement it undergoes in the transition from the reference configuration to the current configuration. Thus, Fig. 2.2.13,[8]

$$\boxed{\mathbf{U}(\mathbf{X},t) = \mathbf{x}(\mathbf{X},t) - \mathbf{X}} \qquad \textbf{Displacement (Material Description)} \qquad (2.2.41)$$

$$\boxed{\mathbf{u}(\mathbf{x},t) = \mathbf{x} - \mathbf{X}(\mathbf{x},t)} \qquad \textbf{Displacement (Spatial Description)} \qquad (2.2.42)$$

Note that **U** and **u** have the same values, they just have different arguments.

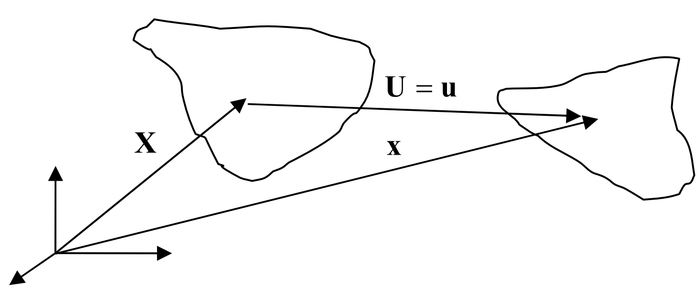


Figure 2.2.13: the displacement

---

[7] In solid mechanics, the motion and deformation are often described in terms of the displacement **u**. In fluid mechanics, however, the primary field quantity describing the kinematic properties is the velocity **v** (and the acceleration $\mathbf{a} = \dot{\mathbf{v}}$) – see later.
[8] The material displacement **U** here is not to be confused with the right stretch tensor discussed earlier.

Displacement Gradients

The displacement gradient in the material and spatial descriptions, $\partial\mathbf{U}(\mathbf{X},t)/\partial\mathbf{X}$ and $\partial\mathbf{u}(\mathbf{x},t)/\partial\mathbf{x}$, are related to the deformation gradient and the inverse deformation gradient through

$$\begin{aligned} \mathrm{Grad}\mathbf{U} &= \frac{\partial\mathbf{U}}{\partial\mathbf{X}} = \frac{\partial(\mathbf{x}-\mathbf{X})}{\partial\mathbf{X}} = \mathbf{F}-\mathbf{I} & \qquad \frac{\partial U_i}{\partial X_j} &= \frac{\partial x_i}{\partial X_j} - \delta_{ij} \\ \mathrm{grad}\mathbf{u} &= \frac{\partial\mathbf{u}}{\partial\mathbf{x}} = \frac{\partial(\mathbf{x}-\mathbf{X})}{\partial\mathbf{x}} = \mathbf{I}-\mathbf{F}^{-1} & \qquad \frac{\partial u_i}{\partial x_j} &= \delta_{ij} - \frac{\partial X_i}{\partial x_j} \end{aligned} \qquad (2.2.43)$$

and it is clear that the displacement gradients are related through (see Eqn. 2.2.8)

$$\mathrm{grad}\mathbf{u} = \mathrm{Grad}\,\mathbf{U}\,\mathbf{F}^{-1} \qquad (2.2.44)$$

The deformation can now be written in terms of either the material or spatial displacement gradients:

$$\begin{aligned} d\mathbf{x} &= d\mathbf{X} + d\mathbf{U}(\mathbf{X}) = d\mathbf{X} + \mathrm{Grad}\mathbf{U}\,d\mathbf{X} \\ d\mathbf{x} &= d\mathbf{X} + d\mathbf{u}(\mathbf{x}) = d\mathbf{X} + \mathrm{grad}\mathbf{u}\,d\mathbf{x} \end{aligned} \qquad (2.2.45)$$

Example

Consider again the extension of the bar shown in Fig. 2.1.5. The displacement is

$$\mathbf{U}(\mathbf{X}) = \left(t + 3X_1 t\right)\mathbf{E}_1, \qquad \mathbf{u}(\mathbf{x}) = \left(\frac{t+3x_1 t}{1+3t}\right)\mathbf{e}_1$$

and the displacement gradients are

$$\mathrm{Grad}\mathbf{U} = 3t\mathbf{E}_1, \qquad \mathrm{grad}\mathbf{u} = \left(\frac{3t}{1+3t}\right)\mathbf{e}_1$$

The displacement is plotted in Fig. 2.2.14 for $t=1$. The two gradients $\partial U_1/\partial X_1$ and $\partial u_1/\partial x_1$ have different values (see the horizontal axes on Fig. 2.2.14). In this example, $\partial U_1/\partial X_1 > \partial u_1/\partial x_1$ – the change in displacement is not as large when "seen" from the spatial coordinates.

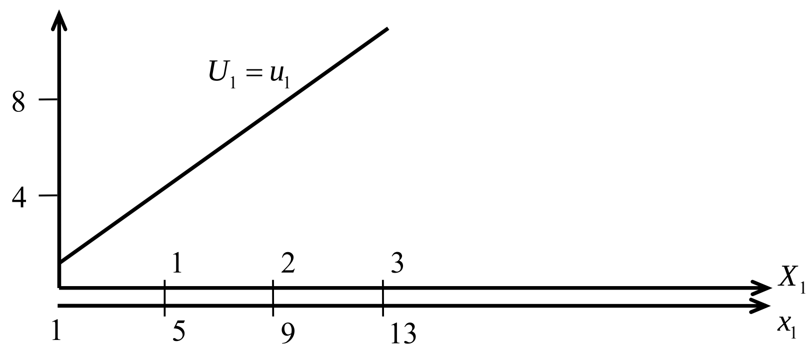


Figure 2.1.14: displacement and displacement gradient

■

Strains in terms of Displacement Gradients

The strains can be written in terms of the displacement gradients. Using 1.10.3b,

$$\begin{aligned}\mathbf{E} &= \frac{1}{2}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F} - \mathbf{I}\right) \\ &= \frac{1}{2}\left((\mathrm{Grad}\mathbf{U} + \mathbf{I})^{\mathrm{T}}(\mathrm{Grad}\mathbf{U} + \mathbf{I}) - \mathbf{I}\right) \\ &= \frac{1}{2}\left(\mathrm{Grad}\mathbf{U} + (\mathrm{Grad}\mathbf{U})^{\mathrm{T}} + (\mathrm{Grad}\mathbf{U})^{\mathrm{T}}\mathrm{Grad}\mathbf{U}\right), \quad E_{IJ} = \frac{1}{2}\left\{\frac{\partial U_I}{\partial X_J} + \frac{\partial U_J}{\partial X_I} + \frac{\partial U_K}{\partial X_I}\frac{\partial U_K}{\partial X_J}\right\}\end{aligned} \tag{2.2.46a}$$

$$\begin{aligned}\mathbf{e} &= \frac{1}{2}\left(\mathbf{I} - \mathbf{F}^{-\mathrm{T}}\mathbf{F}^{-1}\right) \\ &= \frac{1}{2}\left(\mathbf{I} - (\mathbf{I} - \mathrm{grad}\mathbf{u})^{\mathrm{T}}(\mathbf{I} - \mathrm{grad}\mathbf{u})\right) \\ &= \frac{1}{2}\left(\mathrm{grad}\mathbf{u} + (\mathrm{grad}\mathbf{u})^{\mathrm{T}} - (\mathrm{grad}\mathbf{u})^{\mathrm{T}}\mathrm{grad}\mathbf{u}\right), \quad e_{ij} = \frac{1}{2}\left\{\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{\partial u_k}{\partial x_i}\frac{\partial u_k}{\partial x_j}\right\}\end{aligned} \tag{2.2.46b}$$

Small Strain

If the displacement gradients are small, then the quadratic terms, their products, are small relative to the gradients themselves, and may be neglected. With this assumption, the Green-Lagrange strain **E** (and the Euler-Almansi strain) reduces to the **small-strain tensor**,

$$\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathrm{Grad}\mathbf{U} + (\mathrm{Grad}\mathbf{U})^{\mathrm{T}}\right), \quad \varepsilon_{IJ} = \frac{1}{2}\left(\frac{\partial U_I}{\partial X_J} + \frac{\partial U_J}{\partial X_I}\right) \tag{2.2.47}$$

Since in this case the displacement gradients are small, it does not matter whether one refers the strains to the reference or current configurations – the error is of the same order as the quadratic terms already neglected[9], so the small strain tensor can equally well be written as

$$\boxed{\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathrm{grad}\mathbf{u} + (\mathrm{grad}\mathbf{u})^{\mathrm{T}}\right), \qquad \varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)} \qquad \textbf{Small Strain Tensor} \qquad (2.2.48)$$

2.2.8 The Deformation of Area and Volume Elements

Line elements transform between the reference and current configurations through the deformation gradient. Here, the transformation of area and volume elements is examined.

The Jacobian Determinant

The **Jacobian determinant** of the deformation is defined as the determinant of the deformation gradient,

$$\boxed{J(\mathbf{X},t) = \det \mathbf{F}} \qquad \det\mathbf{F} = \begin{vmatrix} \dfrac{\partial x_1}{\partial X_1} & \dfrac{\partial x_1}{\partial X_2} & \dfrac{\partial x_1}{\partial X_3} \\ \dfrac{\partial x_2}{\partial X_1} & \dfrac{\partial x_2}{\partial X_2} & \dfrac{\partial x_2}{\partial X_3} \\ \dfrac{\partial x_3}{\partial X_1} & \dfrac{\partial x_3}{\partial X_2} & \dfrac{\partial x_3}{\partial X_3} \end{vmatrix} \qquad \textbf{The Jacobian Determinant} \qquad (2.2.49)$$

Equivalently, it can be considered to be the Jacobian of the transformation from material to spatial coordinates (see Appendix 1.B.2).

From Eqn. 1.3.17, the Jacobian can also be written in the form of the triple scalar product

$$J = \frac{\partial \mathbf{x}}{\partial X_1} \cdot \left(\frac{\partial \mathbf{x}}{\partial X_2} \times \frac{\partial \mathbf{x}}{\partial X_3}\right) \qquad (2.2.50)$$

Consider now a volume element in the reference configuration, a parallelepiped bounded by the three line-elements $d\mathbf{X}^{(1)}$, $d\mathbf{X}^{(2)}$ and $d\mathbf{X}^{(3)}$. The volume of the parallelepiped[10] is given by the triple scalar product (Eqns. 1.1.4):

$$dV = d\mathbf{X}^{(1)} \cdot \left(d\mathbf{X}^{(2)} \times d\mathbf{X}^{(3)}\right) \qquad (2.2.51)$$

After deformation, the volume element is bounded by the three vectors $d\mathbf{x}^{(i)}$, so that the volume of the deformed element is, using 1.10.16f,

[9] although large rigid body rotations must not be allowed – see §2.7 .
[10] the vectors should form a right-handed set so that the volume is positive.

$$\begin{aligned} dv &= d\mathbf{x}^{(1)} \cdot \left(d\mathbf{x}^{(2)} \times d\mathbf{x}^{(3)} \right) \\ &= \mathbf{F}d\mathbf{X}^{(1)} \cdot \left(\mathbf{F}d\mathbf{X}^{(2)} \times \mathbf{F}d\mathbf{X}^{(3)} \right) \\ &= \det \mathbf{F} \left(d\mathbf{X}^{(1)} \cdot d\mathbf{X}^{(2)} \times d\mathbf{X}^{(3)} \right) \\ &= \det \mathbf{F}\ dV \end{aligned} \tag{2.2.52}$$

Thus the scalar J is a measure of how the volume of a material element has changed with the deformation and for this reason is often called the **volume ratio**.

$$\boxed{dv = J\, dV} \qquad \textbf{Volume Ratio} \tag{2.2.53}$$

Since volumes cannot be negative, one must insist on physical grounds that $J > 0$. Also, since $\mathbf{F}$ has an inverse, $J \neq 0$. Thus one has the restriction

$$J > 0 \tag{2.2.54}$$

Note that a rigid body rotation does not alter the volume, so the volume change is completely characterised by the stretching tensor $\mathbf{U}$. Three line elements lying along the principal directions of $\mathbf{U}$ form an element with volume dV, and then undergo pure stretch into new line elements defining an element of volume $dv = \lambda_1 \lambda_2 \lambda_3 dV$, where λ_i are the principal stretches, Fig. 2.2.15. The unit change in volume is therefore also

$$\frac{dv - dV}{dV} = \lambda_1 \lambda_2 \lambda_3 - 1 \tag{2.2.55}$$

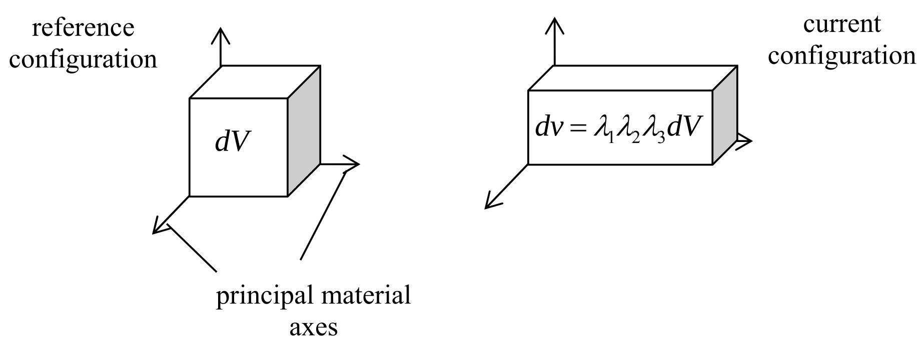


Figure 2.2.15: change in volume

For example, the volume change for pure shear is $-k^2$ (volume decreasing) and, for simple shear, is zero (*cf*. Eqn. 2.2.40 *et seq*., $(\sec\theta + \tan\theta)(\sec\theta - \tan\theta)(1) - 1 = 0$).

An **incompressible** material is one for which the volume change is zero, i.e. the deformation is isochoric. For such a material, $J = 1$, and the three principal stretches are not independent, but are constrained by

$$\boxed{\lambda_1 \lambda_2 \lambda_3 = 1} \qquad \textbf{Incompressibility Constraint} \tag{2.2.56}$$

Nanson's Formula

Consider an area element in the reference configuration, with area dS, unit normal $\hat{\mathbf{N}}$, and bounded by the vectors $d\mathbf{X}^{(1)}, d\mathbf{X}^{(2)}$, Fig. 2.2.16. Then

$$\hat{\mathbf{N}}dS = d\mathbf{X}^{(1)} \times d\mathbf{X}^{(2)} \tag{2.2.57}$$

The volume of the element bounded by the vectors $d\mathbf{X}^{(1)}$, $d\mathbf{X}^{(2)}$ and some arbitrary line element $d\mathbf{X}$ is $dV = \hat{\mathbf{N}}dS \cdot d\mathbf{X}$. The area element is now deformed into an element of area ds with normal $\hat{\mathbf{n}}$ and bounded by the line elements $d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}$. The volume of the new element bounded by the area element and $d\mathbf{x} = \mathbf{F}d\mathbf{X}$ is then

$$dv = \hat{\mathbf{n}}ds \cdot d\mathbf{x} = \hat{\mathbf{n}}ds \cdot \mathbf{F}d\mathbf{X} \equiv J\hat{\mathbf{N}}dS \cdot d\mathbf{X} \tag{2.2.58}$$

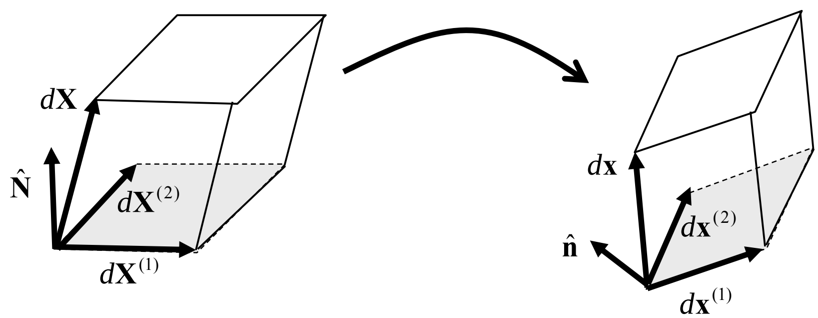


Figure 2.2.16: change of surface area

Thus, since $d\mathbf{X}$ is arbitrary, and using 1.10.3d,

$$\boxed{\hat{\mathbf{n}}ds = J\,\mathbf{F}^{-\mathrm{T}}\hat{\mathbf{N}}dS} \qquad \textbf{Nanson's Formula} \tag{2.2.59}$$

Nanson's formula shows how the vector element of area $\hat{\mathbf{n}}ds$ in the current configuration is related to the vector element of area $\hat{\mathbf{N}}dS$ in the reference configuration.

2.2.9 Inextensibility and Orientation Constraints

A constraint on the principal stretches was introduced for an incompressible material, 2.2.56. Other constraints arise in practice. For example, consider a material which is inextensible in a certain direction, defined by a unit vector $\hat{\mathbf{A}}$ in the reference configuration. It follows that $\left|\mathbf{F}\hat{\mathbf{A}}\right| = 1$ and the constraint can be expressed as 2.2.17,

$$\boxed{\hat{\mathbf{A}}\mathbf{C}\hat{\mathbf{A}} = 1} \qquad \textbf{Inextensibility Constraint} \tag{2.2.60}$$

If there are two such directions in a plane, defined by $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$, making angles θ and ϕ respectively with the principal material axes $\hat{\mathbf{N}}_1, \hat{\mathbf{N}}_2$, then

$$1 = \begin{bmatrix}\cos\theta & \sin\theta & 0\end{bmatrix}\begin{bmatrix}\lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2\end{bmatrix}\begin{bmatrix}\cos\theta \\ \sin\theta \\ 0\end{bmatrix}$$

and $\left(\lambda_1^2 - \lambda_2^2\right)\cos^2\theta = 1 - \lambda_2^2 = \left(\lambda_1^2 - \lambda_2^2\right)\cos^2\phi$. It follows that $\phi = \theta$, $\phi = \theta + \pi$, $\theta + \phi = \pi$ or $\theta + \phi = 2\pi$ (or $\lambda_1 = \lambda_2 = 1$, i.e. no deformation).

Similarly, one can have orientation constraints. For example, suppose that the direction associated with the vector $\hat{\mathbf{A}}$ maintains that direction. Then

$$\boxed{\mathbf{F}\hat{\mathbf{A}} = \mu\hat{\mathbf{A}}} \qquad \textbf{Orientation Constraint} \qquad (2.2.61)$$

for some scalar $\mu > 0$.

2.2.10 Problems

1. In equations 2.2.8, one has from the chain rule
$$\operatorname{grad}\phi = \frac{\partial\phi}{\partial x_i}\mathbf{e}_i = \frac{\partial\phi}{\partial X_m}\frac{\partial X_m}{\partial x_i}\mathbf{e}_i = \left(\frac{\partial\phi}{\partial X_j}\mathbf{E}_j\right)\left(\frac{\partial X_m}{\partial x_i}\mathbf{E}_m \otimes \mathbf{e}_i\right) = \operatorname{Grad}\phi\,\mathbf{F}^{-1}$$
Derive the other two relations.
2. Take the dot product $(\lambda d\hat{\mathbf{x}})\cdot(\lambda d\hat{\mathbf{x}})$ in Eqn. 2.2.29. Then use $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$, $\mathbf{U}^{\mathrm{T}} = \mathbf{U}$, and 1.10.3e to show that
$$\lambda^2 = \frac{d\mathbf{X}}{|d\mathbf{X}|}\mathbf{U}\cdot\mathbf{U}\frac{d\mathbf{X}}{|d\mathbf{X}|}$$
3. For the deformation
$$x_1 = X_1 + 2X_3, \quad x_2 = X_2 - 2X_3, \quad x_3 = -2X_1 + 2X_2 + X_3$$
   (a) Determine the Deformation Gradient and the Right Cauchy-Green tensors
   (b) Consider the two line elements $d\mathbf{X}^{(1)} = \mathbf{e}_1$, $d\mathbf{X}^{(2)} = \mathbf{e}_2$ (emanating from (0,0,0)). Use the Right Cauchy Green tensor to determine whether these elements in the current configuration ($d\mathbf{x}^{(1)}$, $d\mathbf{x}^{(2)}$) are perpendicular.
   (c) Use the right Cauchy Green tensor to evaluate the stretch of the line element $d\mathbf{X} = \mathbf{e}_1 + \mathbf{e}_2$, and hence determine whether the element contracts, stretches, or stays the same length after deformation.
   (d) Determine the Green-Lagrange and Eulerian strain tensors
   (e) Decompose the deformation into a stretching and rotation (check that $\mathbf{U}$ is symmetric and $\mathbf{R}$ is orthogonal). What are the principal stretches?
4. Derive Equations 2.2.36.
5. For the deformation
$$x_1 = X_1, \quad x_2 = X_2 + X_3, \quad x_3 = aX_2 + X_3$$

(a) Determine the displacement vector in both the material and spatial forms
(b) Determine the displaced location of the particles in the undeformed state which originally comprise
   (i) the plane circular surface $X_1 = 0, \quad X_2^2 + X_3^2 = 1/(1-a^2)$
   (ii) the infinitesimal cube with edges along the coordinate axes of length $dX_i = \varepsilon$

   Sketch the displaced configurations if $a = 1/2$

6. For the deformation
$$x_1 = X_1 + aX_2, \quad x_2 = X_2 + aX_3, \quad x_3 = aX_1 + X_3$$
(a) Determine the displacement vector in both the material and spatial forms
(b) Calculate the full material (Green-Lagrange) strain tensor and the full spatial strain tensor
(c) Calculate the infinitesimal strain tensor as derived from the material and spatial tensors, and compare them for the case of very small *a*.

7. In the example given above on the polar decomposition, §2.2.5, check that the relations $\mathbf{Cn}_i = \lambda \mathbf{n}_i, i = 1,2,3$ are satisfied (with respect to the original axes). Check also that the relations $\mathbf{Cn}'_i = \lambda \mathbf{n}'_i, i = 1,2,3$ are satisfied (here, the eigenvectors are the unit vectors in the second coordinate system, the principal directions of **C**, and **C** is with respect to these axes, i.e. it is diagonal).

2.3 Deformation and Strain: Further Topics

2.3.1 Volumetric and Isochoric Deformations

When analysing materials which are only slightly incompressible, it is useful to decompose the deformation gradient multiplicatively, according to

$$\mathbf{F} = \left(J^{1/3}\mathbf{I}\right)\overline{\mathbf{F}} = J^{1/3}\overline{\mathbf{F}} \tag{2.3.1}$$

From this definition {▲Problem 1},

$$\det \overline{\mathbf{F}} = 1 \tag{2.3.2}$$

and so $\overline{\mathbf{F}}$ characterises a volume preserving (**distortional** or **isochoric**) deformation. The tensor $J^{1/3}\mathbf{I}$ characterises the volume-changing (**dilational** or **volumetric**) component of the deformation, with $\det\left(J^{1/3}\mathbf{I}\right) = \det \mathbf{F} = J$.

This concept can be carried on to other kinematic tensors. For example, with $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$,

$$\mathbf{C} = J^{2/3}\overline{\mathbf{F}}^{\mathrm{T}}\overline{\mathbf{F}} \equiv J^{2/3}\overline{\mathbf{C}}. \tag{2.3.3}$$

$\overline{\mathbf{F}}$ and $\overline{\mathbf{C}}$ are called the **modified deformation gradient** and the **modified right Cauchy-Green tensor**, respectively. The square of the stretch is given by

$$\lambda^2 = d\hat{\mathbf{X}}\mathbf{C}d\hat{\mathbf{X}} = J^{2/3}\left\{d\hat{\mathbf{X}}\overline{\mathbf{C}}d\hat{\mathbf{X}}\right\} \tag{2.3.4}$$

so that $\lambda = J^{1/3}\overline{\lambda}$, where $\overline{\lambda}$ is the **modified stretch**, due to the action of $\overline{\mathbf{C}}$. Similarly, the **modified principal stretches** are

$$\overline{\lambda}_i = J^{-1/3}\lambda_i, \qquad i = 1,2,3 \tag{2.3.5}$$

with

$$\det \overline{\mathbf{F}} = \overline{\lambda}_1\overline{\lambda}_2\overline{\lambda}_3 = 1 \tag{2.3.6}$$

The case of simple shear discussed earlier is an example of an isochoric deformation, in which the deformation gradient and the modified deformation gradient coincide, $J^{1/3}\mathbf{I} = \mathbf{I}$.

2.3.2 Relative Deformation

It is usual to use the configuration at $(\mathbf{X}, t = 0)$ as the reference configuration, and define quantities such as the deformation gradient relative to this reference configuration. As mentioned, any configuration can be taken to be the reference configuration, and a new deformation gradient can be constructed with respect to this new reference configuration. Further, the reference configuration does not have to be fixed, but could be moving also.

In many cases, it is useful to choose the *current* configuration $(\mathbf{x}, t)$ to be the reference configuration, for example when evaluating rates of change of kinematic quantities (see later). To this end, introduce a third configuration: this is the configuration at some time $t = \tau$ and the position of a material particle $\mathbf{X}$ here is denoted by $\hat{\mathbf{x}} = \boldsymbol{\chi}(\mathbf{X}, \tau)$, where $\boldsymbol{\chi}$ is the motion function. The deformation at this time τ relative to the *current* configuration is called the **relative deformation**, and is denoted by $\hat{\mathbf{x}} = \boldsymbol{\chi}_{(t)}(\mathbf{x}, \tau)$, as illustrated in Fig. 2.3.1.

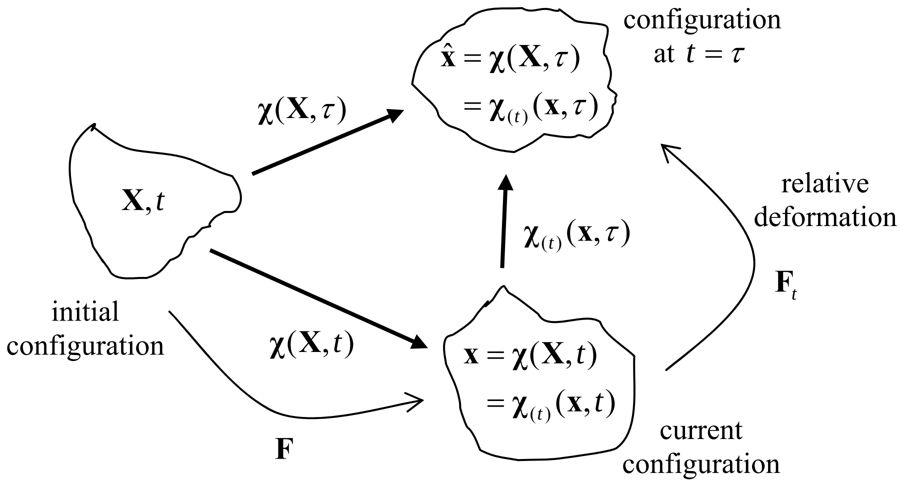


Figure 2.3.1: the relative deformation

The **relative deformation gradient $\mathbf{F}_t$** is defined through

$$d\hat{\mathbf{x}} = \mathbf{F}_t(\mathbf{x}, \tau)\, d\mathbf{x}, \qquad \mathbf{F}_t = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} \tag{2.3.7}$$

Also, since $d\mathbf{x} = \mathbf{F}(\mathbf{X}, t)\, d\mathbf{X}$ and $d\hat{\mathbf{x}} = \mathbf{F}(\mathbf{X}, \tau)\, d\mathbf{X}$, one has the relation

$$\mathbf{F}(\mathbf{X}, \tau) = \mathbf{F}_t(\mathbf{x}, \tau)\, \mathbf{F}(\mathbf{X}, t) \tag{2.3.8}$$

Similarly, relative strain measures can be defined, for example the relative right Cauchy-Green strain tensor is

$$\mathbf{C}_t(\tau) = \mathbf{F}_t(\tau)^{\mathrm{T}} \mathbf{F}_t(\tau) \tag{2.3.9}$$

Example

Consider the two-dimensional motion

$$x_1 = X_1 e^t, \quad x_2 = X_2(t+1)$$

Inverting these gives the spatial description $X_1 = x_1 e^{-t}$, $X_2 = x_2/(t+1)$, and the relative deformation is

$$\begin{aligned}\hat{x}_1(\mathbf{x},\tau) &= X_1 e^{\tau} = x_1 e^{\tau - t} \\ \hat{x}_2(\mathbf{x},\tau) &= X_2(\tau+1) = x_2(\tau+1)/(t+1)\end{aligned}$$

The deformation gradients are

$$\begin{aligned}\mathbf{F}(\mathbf{X},t) &= \frac{\partial x_i}{\partial X_j}\mathbf{e}_i \otimes \mathbf{E}_j = e^t \mathbf{e}_1 \otimes \mathbf{E}_1 + (t+1)\mathbf{e}_2 \otimes \mathbf{E}_2 \\ \mathbf{F}_t(\mathbf{x},\tau) &= \frac{\partial \hat{x}_i}{\partial x_j}\mathbf{e}_i \otimes \mathbf{e}_j = e^{\tau - t}\mathbf{e}_1 \otimes \mathbf{e}_1 + (\tau+1)/(t+1)\mathbf{e}_2 \otimes \mathbf{e}_2\end{aligned}$$

■

2.3.3 Derivatives of the Stretch

In this section, some useful formulae involving the derivatives of the stretches with respect to the Cauchy-Green strain tensors are derived.

Derivatives with respect to b

First, take the stretches to be functions of the left Cauchy-Green strain **b**. Write **b** using the spatial principal directions $\hat{\mathbf{n}}_i$ as a basis, 2.2.37, so that the total differential can be expressed as

$$d\mathbf{b} = \sum_{i=1}^{3} 2\lambda_i d\lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i + \lambda_i^2 \left[d\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i + \hat{\mathbf{n}}_i \otimes d\hat{\mathbf{n}}_i \right] \qquad (2.3.10)$$

Since $\hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_j = \delta_{ij}$, then

$$\hat{\mathbf{n}}_i d\mathbf{b} \hat{\mathbf{n}}_i = 2\lambda_i d\lambda_i + \lambda_i^2 \left[\hat{\mathbf{n}}_i \cdot d\hat{\mathbf{n}}_i + d\hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_i \right] = 2\lambda_i d\lambda_i \quad \text{(no sum over } i\text{)} \qquad (2.3.11)$$

This last follows since the change in a vector of constant length is always orthogonal to the vector itself (as in the curvature analysis of §1.6.2). Using the property $\mathbf{uTv} = \mathbf{T} : (\mathbf{u} \otimes \mathbf{v})$, one has (summing over the *k* but not over the *i*; here $d\lambda_k / d\lambda_i = \delta_{ik}$)

$$d\mathbf{b} : (\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i) \equiv \frac{\partial \mathbf{b}}{\partial \lambda_k} d\lambda_k : (\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i) = 2\lambda_i d\lambda_i \quad \rightarrow \quad \frac{1}{2\lambda_i}\frac{\partial \mathbf{b}}{\partial \lambda_i} : (\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i) = 1 \qquad (2.3.12)$$

Then, since $\partial \mathbf{b} / \partial \lambda_i : \partial \lambda_i / \partial \mathbf{b}$ is also equal to 1, one has

$$\frac{1}{2\lambda_i}\frac{\partial \mathbf{b}}{\partial \lambda_i}:(\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i) = \frac{\partial \mathbf{b}}{\partial \lambda_i}:\frac{\partial \lambda_i}{\partial \mathbf{b}} \quad \rightarrow \quad \frac{\partial \lambda_i}{\partial \mathbf{b}} = \frac{1}{2\lambda_i}(\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i) \tag{2.3.13}$$

The chain rule then gives the second derivative.

The above analysis is for distinct principal stretches. When $\lambda_1 = \lambda_2 = \lambda_3 \equiv \lambda$, then $\mathbf{b} = \lambda^2\mathbf{I}$, $d\mathbf{b} = 2\lambda d\lambda \mathbf{I}$. Also, $d\mathbf{b} = 3(\partial \mathbf{b}/\partial\lambda)d\lambda$, so $3(\partial \mathbf{b}/\partial\lambda) = 2\lambda\mathbf{I}$, or

$$3\frac{\partial \mathbf{b}}{\partial \lambda}:\frac{\partial \lambda}{\partial \mathbf{b}} = 2\lambda\mathbf{I}:\frac{\partial \lambda}{\partial \mathbf{b}} \tag{2.3.14}$$

But $\partial \mathbf{b}/\partial\lambda : \partial\lambda/\partial\mathbf{b} = 1$ and $3 = \mathbf{I}:\mathbf{I}$, and so in this case, $\partial\lambda/\partial\mathbf{b} = \mathbf{I}/2\lambda$.

A similar calculation can be carried out for two equal eigenvalues $\lambda_1 = \lambda_2 = \lambda \neq \lambda_3$. In summary,

$$\boxed{\begin{array}{lll}
\dfrac{\partial \lambda_i}{\partial \mathbf{b}} = \dfrac{1}{2\lambda_i}\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i & (\text{no sum over } i) & \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1 \\
\dfrac{\partial \lambda}{\partial \mathbf{b}} = \dfrac{1}{2\lambda}\left(\hat{\mathbf{n}}_1 \otimes \hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2 \otimes \hat{\mathbf{n}}_2\right) & & \\
\dfrac{\partial \lambda_3}{\partial \mathbf{b}} = \dfrac{1}{2\lambda_3}\left(\hat{\mathbf{n}}_3 \otimes \hat{\mathbf{n}}_3\right) & & \lambda_1 = \lambda_2 = \lambda \neq \lambda_3 \\
\dfrac{\partial \lambda}{\partial \mathbf{b}} = \dfrac{1}{2\lambda}\sum_{i=1}^{3}\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i = \dfrac{1}{2\lambda}\mathbf{I} & & \lambda_1 = \lambda_2 = \lambda_3 = \lambda \\
 & & \\
\dfrac{\partial^2 \lambda_i}{\partial \mathbf{b}^2} = -\dfrac{1}{4\lambda_i^3}\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i & (\text{no sum over } i) & \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1
\end{array}} \tag{2.3.15}$$

Derivatives with respect to C

The stretch can also be considered to be a function of the right Cauchy-Green strain **C**. The derivatives of the stretches with respect to **C** can be found in exactly the same way as for the left Cauchy-Green strain. The results are the same as given in 2.3.15 except that, referring to 2.2.37, **b** is replaced by **C** and $\hat{\mathbf{n}}$ is replaced by $\hat{\mathbf{N}}$.

2.3.4 The Directional Derivative of Kinematic Quantities

The directional derivative of vectors and tensors was introduced in §1.6.11 and §1.15.4. Taking directional derivatives of kinematic quantities is often very useful, for example in linearising equations in order to apply numerical solution algorithms

The Deformation Gradient

First, consider the deformation gradient as a function of the current position **x** (or motion $\boldsymbol{\chi}$) and examine its value at $\mathbf{x} + \mathbf{a}$:

$$\mathbf{F}(\mathbf{x}+\mathbf{a}) = \mathbf{F}(\mathbf{x}) + \partial_{\mathbf{x}}\mathbf{F}[\mathbf{a}] + o(|\mathbf{a}|) \tag{2.3.16}$$

The directional derivative $\partial_{\mathbf{x}}\mathbf{F}[\mathbf{a}] = (\partial\mathbf{F}/\partial\mathbf{x})\mathbf{a}$ can be expressed as

$$\begin{aligned}
\partial_{\mathbf{x}}\mathbf{F}[\mathbf{a}] &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{F}(\mathbf{x}+\varepsilon\mathbf{a}) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{\partial(\mathbf{x}+\varepsilon\mathbf{a})}{\partial\mathbf{X}} \\
&= \mathrm{Grad}\,\mathbf{a} \\
&= (\mathrm{grad}\,\mathbf{a})\mathbf{F}
\end{aligned} \tag{2.3.17}$$

the last line resulting from 2.2.8b. It follows that the directional derivative of the deformation gradient in the direction of a displacement vector **u** from the *current* configuration is

$$\partial_{\mathbf{x}}\mathbf{F}[\mathbf{u}] = (\mathrm{grad}\,\mathbf{u})\mathbf{F} \tag{2.3.18}$$

On the other hand, consider the deformation gradient as a function of **X** and examine its value at $\mathbf{X}+\mathbf{A}$:

$$\mathbf{F}(\mathbf{X}+\mathbf{A}) = \mathbf{F}(\mathbf{X}) + \partial_{\mathbf{X}}\mathbf{F}[\mathbf{A}] \tag{2.3.19}$$

and now

$$\begin{aligned}
\partial_{\mathbf{X}}\mathbf{F}[\mathbf{A}] &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{F}(\mathbf{X}+\varepsilon\mathbf{A}) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{\partial}{\partial\mathbf{X}}\mathbf{x}(\mathbf{X}+\varepsilon\mathbf{A}) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{\partial}{\partial\mathbf{X}}(\mathbf{x}+\mathbf{F}\varepsilon\mathbf{A}) \\
&= \mathrm{Grad}(\mathbf{F}\mathbf{A}) \\
&= \mathrm{Grad}\,\mathbf{a}
\end{aligned} \tag{2.3.20}$$

where $\mathbf{a} = \mathbf{F}\mathbf{A}$.

Other Kinematic Quantities

The directional derivative of the Green-Lagrange strain, the right and left Cauchy-Green tensors and the Jacobian in the direction of a displacement **u** from the current configuration are {▲Problem 2}

$$\begin{aligned}
\partial_{\mathbf{x}}\mathbf{E}[\mathbf{u}] &= \mathbf{F}^{\mathrm{T}}\boldsymbol{\varepsilon}\mathbf{F} \\
\partial_{\mathbf{x}}\mathbf{C}[\mathbf{u}] &= 2\mathbf{F}^{\mathrm{T}}\boldsymbol{\varepsilon}\mathbf{F} \\
\partial_{\mathbf{x}}\mathbf{b}[\mathbf{u}] &= (\mathrm{grad}\mathbf{u})\mathbf{b} + \mathbf{b}(\mathrm{grad}\mathbf{u})^{\mathrm{T}} \\
\partial_{\mathbf{x}}J[\mathbf{u}] &= J\mathrm{div}\mathbf{u}
\end{aligned} \tag{2.3.21}$$

where $\boldsymbol{\varepsilon}$ is the small-strain tensor, 2.2.48.

The directional derivative is also useful for deriving various relations between the kinematic variables. For example, for an arbitrary vector **a**, using the chain rule 1.15.28, 2.3.20, 1.15.24, the trace relations 1.10.10e and 1.10.10b, and 2.2.8b, 1.14.9,

$$\begin{aligned}
(\mathrm{Grad}J)\cdot\mathbf{a} &= \partial_{\mathbf{X}}J[\mathbf{a}] \\
&= \partial_{\mathbf{F}}J[\partial_{\mathbf{X}}\mathbf{F}[\mathbf{a}]] \\
&= \partial_{\mathbf{F}}J[\mathrm{Grad}(\mathbf{Fa})] \\
&= J\mathbf{F}^{-\mathrm{T}} : \mathrm{Grad}(\mathbf{Fa}) \\
&= J\mathrm{tr}\left(\mathbf{F}^{-1}\mathrm{Grad}(\mathbf{Fa})\right) \\
&= J\mathrm{tr}\left(\mathrm{Grad}(\mathbf{Fa})\mathbf{F}^{-1}\right) \\
&= J\mathrm{tr}(\mathrm{grad}(\mathbf{Fa})) \\
&= J\,\mathrm{div}(\mathbf{Fa})
\end{aligned} \tag{2.3.22}$$

so that, from 1.14.16b with **a** constant,

$$\boxed{\mathrm{Grad}J = J\mathrm{div}\mathbf{F}^{\mathrm{T}}} \tag{2.3.23}$$

2.3.5 Problems

1. Use 1.10.16c to show that $\det\overline{\mathbf{F}} = 1$.
2. (a) use the relation $\mathbf{E} = \frac{1}{2}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F} - \mathbf{I}\right)$, Eqn. 2.3.18, $\partial_{\mathbf{x}}\mathbf{F}[\mathbf{u}] = (\mathrm{grad}\mathbf{u})\mathbf{F}$, and the product rule of differentiation to derive 2.3.21a, $\partial_{\mathbf{x}}\mathbf{E}[\mathbf{u}] = \mathbf{F}^{\mathrm{T}}\boldsymbol{\varepsilon}\mathbf{F}$, where $\boldsymbol{\varepsilon}$ is the small strain tensor.
   (b) evaluate $\partial_{\mathbf{x}}\mathbf{C}[\mathbf{u}]$ (in terms of **F** and $\boldsymbol{\varepsilon}$, the small strain tensor)
   (c) evaluate $\partial_{\mathbf{x}}\mathbf{b}[\mathbf{u}]$ (in terms of $\mathrm{grad}\mathbf{u}$ and **b**)
   (d) evaluate $\partial_{\mathbf{x}}J[\mathbf{u}]$ (in terms of J and $\mathrm{div}\mathbf{u}$; use the chain rule $\partial_{\mathbf{x}}J[\mathbf{u}] = \partial_{\mathbf{F}}\hat{J}[\partial_{\mathbf{x}}\mathbf{F}[\mathbf{u}]]$, with $\hat{J}(\mathbf{F}) = \det\mathbf{F}$, $\partial_{\mathbf{x}}\mathbf{F}[\mathbf{u}] = \mathrm{Grad}\mathbf{u}$)

2.4 Material Time Derivatives

The motion is now allowed to be a function of time, $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$, and attention is given to time derivatives, both the **material time derivative** and the **local time derivative**.

2.4.1 Velocity & Acceleration

The velocity of a moving particle is the time rate of change of the position of the particle. From 2.1.3, by definition,

$$\mathbf{V}(\mathbf{X}, t) \equiv \frac{d\boldsymbol{\chi}(\mathbf{X}, t)}{dt} \tag{2.4.1}$$

In the motion expression $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$, $\mathbf{X}$ and t are independent variables and $\mathbf{X}$ is independent of time, denoting the particle for which the velocity is being calculated. The velocity can thus be written as $\partial \boldsymbol{\chi}(\mathbf{X}, t)/\partial t$ or, denoting the motion by $\mathbf{x}(\mathbf{X}, t)$, as $d\mathbf{x}(\mathbf{X}, t)/dt$ or $\partial \mathbf{x}(\mathbf{X}, t)/\partial t$.

The spatial description of the velocity field may be obtained from the material description by simply replacing $\mathbf{X}$ with $\mathbf{x}$, i.e.

$$\mathbf{v}(\mathbf{x}, t) = \mathbf{V}\left(\boldsymbol{\chi}^{-1}(\mathbf{x}, t), t\right) \tag{2.4.2}$$

As with displacements in both descriptions, there is only *one* velocity, $\mathbf{V}(\mathbf{X}, t) = \mathbf{v}(\mathbf{x}, t)$ – they are just given in terms of different coordinates.

The velocity is most often expressed in the spatial description, as

$$\boxed{\mathbf{v}(\mathbf{x}, t) = \dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}} \qquad \textbf{velocity} \tag{2.4.3}$$

To be precise, the right hand side here involves $\mathbf{x}$ which is a function of the material coordinates, but it is understood that the substitution back to spatial coordinates, as in 2.4.2, is made (see example below).

Similarly, the acceleration is defined to be

$$\mathbf{A}(\mathbf{X}, t) = \frac{d^2\boldsymbol{\chi}(\mathbf{X}, t)}{dt^2} = \frac{d^2\mathbf{x}}{dt^2} = \frac{d\mathbf{V}}{dt} = \frac{\partial^2\boldsymbol{\chi}(\mathbf{X}, t)}{\partial t^2} \tag{2.4.4}$$

Example

Consider the motion

$$x_1 = X_1 + t^2 X_2, \quad x_2 = X_2 + t^2 X_1, \quad x_3 = X_3$$

The velocity and acceleration can be evaluated through

$$\mathbf{V}(\mathbf{X},t) = \frac{d\mathbf{x}}{dt} = 2tX_2\mathbf{e}_1 + 2tX_1\mathbf{e}_2, \qquad \mathbf{A}(\mathbf{X},t) = \frac{d^2\mathbf{x}}{dt^2} = 2X_2\mathbf{e}_1 + 2X_1\mathbf{e}_2$$

One can write the motion in the spatial description by inverting the material description:

$$X_1 = \frac{x_1 - t^2 x_2}{1 - t^4}, \quad X_2 = \frac{x_2 - t^2 x_1}{1 - t^4}, \quad X_3 = x_3$$

Substituting in these equations then gives the spatial description of the velocity and acceleration:

$$\begin{aligned}
\mathbf{v}(\mathbf{x},t) &= \mathbf{V}\left(\chi^{-1}(\mathbf{x},t),t\right) = 2t\frac{x_2 - t^2 x_1}{1-t^4}\mathbf{e}_1 + 2t\frac{x_1 - t^2 x_2}{1-t^4}\mathbf{e}_2 \\
\mathbf{a}(\mathbf{x},t) &= \mathbf{A}\left(\chi^{-1}(\mathbf{x},t),t\right) = 2\frac{x_2 - t^2 x_1}{1-t^4}\mathbf{e}_1 + 2\frac{x_1 - t^2 x_2}{1-t^4}\mathbf{e}_2
\end{aligned}$$

■

2.4.2 The Material Derivative

One can analyse deformation by examining the current configuration only, discounting the reference configuration. This is the viewpoint taken in Fluid Mechanics – one focuses on material as it flows at the *current time*, and does not consider "where the fluid was". In order to do this, quantities must be cast in terms of the velocity. Suppose that the velocity in terms of spatial coordinates, $\mathbf{v} = \mathbf{v}(\mathbf{x},t)$ is known; for example, one could have a measuring instrument which records the velocity at a specific location, but the motion $\boldsymbol{\chi}$ itself is unknown. In that case, to evaluate the acceleration, the chain rule of differentiation must be applied:

$$\dot{\mathbf{v}} \equiv \frac{d}{dt}\mathbf{v}(\mathbf{x}(t),t) = \frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\frac{d\mathbf{x}}{dt}$$

or

$$\boxed{\mathbf{a} = \frac{\partial \mathbf{v}}{\partial t} + (\operatorname{grad}\mathbf{v})\mathbf{v}} \qquad \textbf{acceleration (spatial description)} \qquad (2.4.5)$$

The acceleration can now be determined, because the derivatives can be determined (measured) without knowing the motion.

In the above, the **material derivative**, or **total derivative**, of the particle's velocity was taken to obtain the acceleration. In general, one can take the time derivative of any physical or kinematic property $(\bullet)$ expressed in the spatial description:

$$\boxed{\frac{d}{dt}(\bullet) = \frac{\partial}{\partial t}(\bullet) + \operatorname{grad}(\bullet)\mathbf{v}} \qquad \textbf{Material Time Derivative} \tag{2.4.6}$$

For example, the rate of change of the density $\rho = \rho(\mathbf{x}, t)$ of a particle instantaneously at $\mathbf{x}$ is

$$\dot{\rho} \equiv \frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \operatorname{grad}\rho \cdot \mathbf{v} \tag{2.4.7}$$

The Local Rate of Change

The first term, $\partial\rho / \partial t$, gives the **local rate of change** of density at **x** whereas the second term $\mathbf{v} \cdot \operatorname{grad}\rho$ gives the change due to the particle's motion, and is called the **convective rate of change**.

Note the difference between the material derivative and the local derivative. For example, the material derivative of the velocity, 2.4.5 (or, equivalently, $d\mathbf{V}(\mathbf{X}, t)/dt$ in 2.4.4, with **X** fixed) is not the same as the derivative $\partial\mathbf{v}(\mathbf{x}, t)/\partial t$ (with **x** fixed). The former is the acceleration of a material particle **X**. The latter is the time rate of change of the velocity of particles *at a fixed location* in space; in general, *different* material particles will occupy position **x** at different times.

The material derivative d/dt can be applied to any scalar, vector or tensor:

$$\begin{aligned}
\dot{\alpha} &\equiv \frac{d\alpha}{dt} = \frac{\partial \alpha}{\partial t} + \operatorname{grad}\alpha \cdot \mathbf{v} \\
\dot{\mathbf{a}} &\equiv \frac{d\mathbf{a}}{dt} = \frac{\partial \mathbf{a}}{\partial t} + (\operatorname{grad}\mathbf{a})\mathbf{v} \\
\dot{\mathbf{A}} &\equiv \frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + (\operatorname{grad}\mathbf{A})\mathbf{v}
\end{aligned} \tag{2.4.8}$$

Another notation often used for the material derivative is D/Dt:

$$\frac{Df}{Dt} \equiv \frac{df}{dt} \equiv \dot{f} \tag{2.4.9}$$

Steady and Uniform Flows

In a **steady flow**, quantities are independent of time, so the local rate of change is zero and, for example, $\dot{\rho} = \operatorname{grad}\rho \cdot \mathbf{v}$. In a **uniform flow**, quantities are independent of position so that, for example, $\dot{\rho} = \partial\rho / \partial t$

Example

Consider again the previous example. This time, with only the velocity $\mathbf{v}(\mathbf{x}, t)$ known, the acceleration can be obtained through the material derivative:

$$\mathbf{a}(\mathbf{x},t) = \frac{\partial \mathbf{v}}{\partial t} + (\operatorname{grad}\mathbf{v})\mathbf{v}$$

$$= \frac{\partial}{\partial t}\left(2t\frac{x_2 - t^2 x_1}{1-t^4}\mathbf{e}_1 + 2t\frac{x_1 - t^2 x_2}{1-t^4}\mathbf{e}_2\right) + \begin{bmatrix} -\frac{2t^3}{1-t^4} & \frac{2t}{1-t^4} & 0 \\ \frac{2t}{1-t^4} & -\frac{2t^3}{1-t^4} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 2t\frac{x_2 - t^2 x_1}{1-t^4} \\ 2t\frac{x_1 - t^2 x_2}{1-t^4} \\ 0 \end{bmatrix}$$

$$= 2\frac{x_2 - t^2 x_1}{1-t^4}\mathbf{e}_1 + 2\frac{x_1 - t^2 x_2}{1-t^4}\mathbf{e}_2$$

as before.

■

The Relationship between the Displacement and Velocity

The velocity can be derived directly from the displacement 2.2.42:

$$\mathbf{v} = \frac{d\mathbf{x}}{dt} = \frac{d(\mathbf{u}+\mathbf{X})}{dt} = \frac{d\mathbf{u}}{dt}, \tag{2.4.10}$$

or

$$\mathbf{v} = \frac{d\mathbf{u}}{dt} = \frac{\partial \mathbf{u}}{\partial t} + (\operatorname{grad}\mathbf{u})\mathbf{v} \tag{2.4.11}$$

When the displacement field is given in material form one has

$$\mathbf{V} = \frac{d\mathbf{U}}{dt} \tag{2.4.12}$$

2.4.3 Problems

1. The density of a material is given by

$$\rho = \frac{e^{-2t}}{\mathbf{x}\cdot\mathbf{x}}$$

The velocity field is given by

$$v_1 = x_2 + 2x_3, \quad v_2 = x_3 - 2x_1, \quad v_3 = x_1 + 2x_2$$

Determine the time derivative of the density (a) at a certain position $\mathbf{x}$ in space, and (b) of a material particle instantaneously occupying position $\mathbf{x}$.

2.5 Deformation Rates

In this section, rates of change of the deformation tensors introduced earlier, **F**, **C**, **E**, etc., are evaluated, and special tensors used to measure deformation rates are discussed, for example the velocity gradient **l**, the rate of deformation **d** and the spin tensor **w**.

2.5.1 The Velocity Gradient

The **velocity gradient** is used as a measure of the rate at which a material is deforming.

Consider two fixed neighbouring points, $\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$, Fig. 2.5.1. The velocities of the material particles at these points at any given time instant are $\mathbf{v}(\mathbf{x})$ and $\mathbf{v}(\mathbf{x} + d\mathbf{x})$, and

$$\mathbf{v}(\mathbf{x} + d\mathbf{x}) = \mathbf{v}(\mathbf{x}) + \frac{\partial \mathbf{v}}{\partial \mathbf{x}} d\mathbf{x},$$

The relative velocity between the points is

$$d\mathbf{v} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}} d\mathbf{x} \equiv \mathbf{l} d\mathbf{x} \tag{2.5.1}$$

with **l** defined to be the (spatial) velocity gradient,

$$\boxed{\mathbf{l} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \operatorname{grad} \mathbf{v}, \qquad l_{ij} = \frac{\partial v_i}{\partial x_j}} \qquad \textbf{Spatial Velocity Gradient} \tag{2.5.2}$$

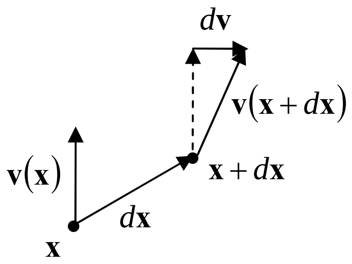


Figure 2.5.1: velocity gradient

Expression 2.5.1 emphasises the tensorial character of the spatial velocity gradient, mapping as it does one vector into another. Its physical meaning will become clear when it is decomposed into its symmetric and skew-symmetric parts below.

The spatial velocity gradient is commonly used in both solid and fluid mechanics. Less commonly used is the material velocity gradient, which is related to the rate of change of the deformation gradient:

$$\operatorname{Grad} \mathbf{V} = \frac{\partial \mathbf{V}(\mathbf{X}, t)}{\partial \mathbf{X}} = \frac{\partial}{\partial \mathbf{X}} \left(\frac{\partial \mathbf{x}(\mathbf{X}, t)}{\partial t} \right) = \frac{\partial}{\partial t} \left(\frac{\partial \mathbf{x}(\mathbf{X}, t)}{\partial \mathbf{X}} \right) = \dot{\mathbf{F}} \tag{2.5.3}$$

and use has been made of the fact that, since **X** and t are independent variables, material time derivatives and material gradients commute.

2.5.2 Material Derivatives of the Deformation Gradient

The spatial velocity gradient may be written as

$$\frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \frac{\partial \mathbf{v}}{\partial \mathbf{X}}\frac{\partial \mathbf{X}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{X}}\left(\frac{\partial \mathbf{x}}{\partial t}\right)\frac{\partial \mathbf{X}}{\partial \mathbf{x}} = \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{X}}\right)\frac{\partial \mathbf{X}}{\partial \mathbf{x}}$$

or $\mathbf{l} = \dot{\mathbf{F}}\mathbf{F}^{-1}$ so that the material derivative of **F** can be expressed as

$$\boxed{\dot{\mathbf{F}} = \mathbf{l}\,\mathbf{F}} \qquad \textbf{Material Time Derivative of the Deformation Gradient} \qquad (2.5.4)$$

Also, it can be shown that {▲Problem 1}

$$\boxed{\begin{aligned} \overline{\mathbf{F}^{\mathrm{T}}}^{\cdot} &= \dot{\mathbf{F}}^{\mathrm{T}} \\ \overline{\mathbf{F}^{-1}}^{\cdot} &= -\mathbf{F}^{-1}\mathbf{l} \\ \overline{\mathbf{F}^{-\mathrm{T}}}^{\cdot} &= -\mathbf{l}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}} \end{aligned}} \qquad (2.5.5)$$

2.5.3 The Rate of Deformation and Spin Tensors

The velocity gradient can be decomposed into a symmetric tensor and a skew-symmetric tensor as follows (see §1.10.10):

$$\boxed{\mathbf{l} = \mathbf{d} + \mathbf{w}} \qquad (2.5.6)$$

where **d** is the **rate of deformation tensor** (or **rate of stretching tensor**) and **w** is the **spin tensor** (or **rate of rotation**, or **vorticity tensor**), defined by

$$\boxed{\begin{aligned} \mathbf{d} &= \frac{1}{2}\left(\mathbf{l} + \mathbf{l}^{\mathrm{T}}\right), & d_{ij} &= \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \\ \mathbf{w} &= \frac{1}{2}\left(\mathbf{l} - \mathbf{l}^{\mathrm{T}}\right), & w_{ij} &= \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i}\right) \end{aligned}} \qquad \textbf{Rate of Deformation and Spin Tensors}$$

(2.5.7)

The physical meaning of these tensors is next examined.

The Rate of Deformation

Consider first the rate of deformation tensor **d** and note that

$$\mathbf{l}d\mathbf{x} = d\mathbf{v} = \frac{d}{dt}(d\mathbf{x}) \tag{2.5.8}$$

The rate at which the square of the length of $d\mathbf{x}$ is changing is then

$$\begin{aligned}
&\frac{d}{dt}\left(|d\mathbf{x}|^2\right) = 2|d\mathbf{x}|\frac{d}{dt}\left(|d\mathbf{x}|\right), \\
&\frac{d}{dt}\left(|d\mathbf{x}|^2\right) = \frac{d}{dt}\left(d\mathbf{x}\cdot d\mathbf{x}\right) = 2d\mathbf{x}\cdot\frac{d}{dt}\left(d\mathbf{x}\right) = 2d\mathbf{x}\,\mathbf{l}\,d\mathbf{x} = 2d\mathbf{x}\,\mathbf{d}\,d\mathbf{x}
\end{aligned} \tag{2.5.9}$$

the last equality following from 2.5.6 and 1.10.31e. Dividing across by $2|d\mathbf{x}|^2$, then leads to

$$\boxed{\frac{\dot{\lambda}}{\lambda} = \hat{\mathbf{n}}\,\mathbf{d}\hat{\mathbf{n}}} \qquad \textbf{Rate of stretching per unit stretch in the direction } \hat{\mathbf{n}} \tag{2.5.10}$$

where $\lambda = |d\mathbf{x}|/|d\mathbf{X}|$ is the stretch and $\hat{\mathbf{n}} = d\mathbf{x}/|d\mathbf{x}|$ is a unit normal in the direction of $d\mathbf{x}$. Thus the rate of deformation **d** gives the rate of stretching of line elements. The diagonal components of **d**, for example

$$d_{11} = \mathbf{e}_1\mathbf{d}\mathbf{e}_1,$$

represent unit rates of extension in the coordinate directions.

Note that these are *instantaneous* rates of extension, in other words, they are rates of extensions of elements in the current configuration at the current time; they are not a measure of the rate at which a line element in the original configuration changed into the corresponding line element in the current configuration.

Note:

- Eqn. 2.5.10 can also be derived as follows: let $\hat{\mathbf{N}}$ be a unit normal in the direction of $d\mathbf{X}$, and $\hat{\mathbf{n}}$ be the corresponding unit normal in the direction of $d\mathbf{x}$. Then $\hat{\mathbf{n}}|d\mathbf{x}| = \mathbf{F}\hat{\mathbf{N}}|d\mathbf{X}|$, or $\hat{\mathbf{n}}\lambda = \mathbf{F}\hat{\mathbf{N}}$. Differentiating gives $\dot{\hat{\mathbf{n}}}\lambda + \hat{\mathbf{n}}\dot{\lambda} = \dot{\mathbf{F}}\hat{\mathbf{N}} = \mathbf{l}\mathbf{F}\hat{\mathbf{N}}$ or $\dot{\hat{\mathbf{n}}}\lambda + \hat{\mathbf{n}}\dot{\lambda} = \mathbf{l}\hat{\mathbf{n}}\lambda$. Contracting both sides with $\hat{\mathbf{n}}$ leads to $\hat{\mathbf{n}}\cdot\dot{\hat{\mathbf{n}}} + \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}\left(\dot{\lambda}/\lambda\right) = \hat{\mathbf{n}}\mathbf{l}\hat{\mathbf{n}}$. But $\hat{\mathbf{n}}\cdot\hat{\mathbf{n}} = 1 \rightarrow d(\hat{\mathbf{n}}\cdot\hat{\mathbf{n}})dt = 0$ so, by the chain rule, $\hat{\mathbf{n}}\cdot\dot{\hat{\mathbf{n}}} = 0$ (confirming that a vector $\hat{\mathbf{n}}$ of constant length is orthogonal to a change in that vector $d\hat{\mathbf{n}}$), and the result follows

Consider now the rate of change of the angle θ between two vectors $d\mathbf{x}^{(1)}, d\mathbf{x}^{(2)}$. Using 2.5.8 and 1.10.3d,

$$
\begin{aligned}
\frac{d}{dt}\left(d\mathbf{x}^{(1)} \cdot d\mathbf{x}^{(2)}\right) &= \frac{d}{dt}\left(d\mathbf{x}^{(1)}\right) \cdot d\mathbf{x}^{(2)} + d\mathbf{x}^{(1)} \cdot \frac{d}{dt}\left(d\mathbf{x}^{(2)}\right) \\
&= \mathbf{l}d\mathbf{x}^{(1)} \cdot d\mathbf{x}^{(2)} + d\mathbf{x}^{(1)} \cdot \mathbf{l}d\mathbf{x}^{(2)} \\
&= \left(\mathbf{l} + \mathbf{l}^{\mathrm{T}}\right) d\mathbf{x}^{(1)} \cdot d\mathbf{x}^{(2)} \\
&= 2\, d\mathbf{x}^{(1)} \mathbf{d} d\mathbf{x}^{(2)}
\end{aligned}
\tag{2.5.11}
$$

which reduces to 2.5.9 when $d\mathbf{x}^{(1)} = d\mathbf{x}^{(2)}$. An alternative expression for this dot product is

$$
\begin{aligned}
\frac{d}{dt}\left(\left|d\mathbf{x}^{(1)}\right|\left|d\mathbf{x}^{(2)}\right|\cos\theta\right) &= \frac{d}{dt}\left(\left|d\mathbf{x}^{(1)}\right|\right)\left|d\mathbf{x}^{(2)}\right|\cos\theta + \frac{d}{dt}\left(\left|d\mathbf{x}^{(2)}\right|\right)\left|d\mathbf{x}^{(1)}\right|\cos\theta - \sin\theta\,\dot{\theta}\left|d\mathbf{x}^{(1)}\right|\left|d\mathbf{x}^{(2)}\right| \\
&= \left(\frac{\frac{d}{dt}\left(\left|d\mathbf{x}^{(1)}\right|\right)}{\left|d\mathbf{x}^{(1)}\right|}\cos\theta + \frac{\frac{d}{dt}\left(\left|d\mathbf{x}^{(2)}\right|\right)}{\left|d\mathbf{x}^{(2)}\right|}\cos\theta - \sin\theta\,\dot{\theta}\right)\left|d\mathbf{x}^{(1)}\right|\left|d\mathbf{x}^{(2)}\right|
\end{aligned}
\tag{2.5.12}
$$

Equating 2.5.11 and 2.5.12 leads to

$$
2\,\hat{\mathbf{n}}_1\,\mathbf{d}\,\hat{\mathbf{n}}_2 = \left(\frac{\dot{\lambda}_1}{\lambda_1} + \frac{\dot{\lambda}_2}{\lambda_2}\right)\cos\theta - \sin\theta\,\dot{\theta} \tag{2.5.13}
$$

where $\lambda_i = \left|d\mathbf{x}^{(i)}\right| / \left|d\mathbf{X}^{(i)}\right|$ is the stretch and $\hat{\mathbf{n}}_i = d\mathbf{x}^{(i)} / \left|d\mathbf{x}^{(i)}\right|$ is a unit normal in the direction of $d\mathbf{x}^{(i)}$.

It follows from 2.5.13 that the off-diagonal terms of the rate of deformation tensor represent **shear rates**: the rate of change of the right angle between line elements aligned with the coordinate directions. For example, taking the base vectors $\mathbf{e}_1 = \hat{\mathbf{n}}_1$, $\mathbf{e}_2 = \hat{\mathbf{n}}_2$, 2.5.13 reduces to

$$
d_{12} = -\frac{1}{2}\dot{\theta}_{12} \tag{2.5.14}
$$

where θ_{12} is the instantaneous right angle between the axes in the current configuration.

The Spin

Consider now the spin tensor $\mathbf{w}$; since it is skew-symmetric, it can be written in terms of its axial vector $\boldsymbol{\omega}$ (Eqn. 1.10.34), called the **angular velocity vector**:

$$\begin{aligned}\boldsymbol{\omega} &= -w_{23}\mathbf{e}_1 + w_{13}\mathbf{e}_2 - w_{12}\mathbf{e}_3 \\ &= \frac{1}{2}\left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}\right)\mathbf{e}_1 + \frac{1}{2}\left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}\right)\mathbf{e}_2 + \frac{1}{2}\left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right)\mathbf{e}_3 \\ &= \frac{1}{2}\operatorname{curl}\mathbf{v}\end{aligned} \tag{2.5.15}$$

(The vector $2\boldsymbol{\omega}$ is called the **vorticity** (or **spin**) **vector**.) Thus when $\mathbf{d}$ is zero, the motion consists of a rotation about some axis at angular velocity $\omega = |\boldsymbol{\omega}|$ (*cf*. the end of §1.10.11), with $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, $\mathbf{r}$ measured from a point on the axis, and $\mathbf{wr} = \boldsymbol{\omega} \times \mathbf{r} = \mathbf{v}$.

On the other hand, when $\mathbf{l} = \mathbf{d}$, $\mathbf{w} = \mathbf{0}$, one has $\boldsymbol{\omega} = \mathbf{o}$, and the motion is called **irrotational**.

Example (Shear Flow)

Consider a **simple shear flow** in which the velocity profile is "triangular" as shown in Fig. 2.5.2. This type of flow can be generated (at least approximately) in many fluids by confining the fluid between plates a distance h apart, and by sliding the upper plate over the lower one at constant velocity V. If the material particles adjacent to the upper plate have velocity $V\mathbf{e}_1$, then the velocity field is $\mathbf{v} = \dot{\gamma} x_2 \mathbf{e}_1$, where $\dot{\gamma} = V/h$. This is a steady flow ($\partial \mathbf{v}/\partial t = \mathbf{0}$); at any given point, there is no change over time. The velocity gradient is $\mathbf{l} = \dot{\gamma}\mathbf{e}_1 \otimes \mathbf{e}_2$ and the acceleration of material particles is zero: $\mathbf{a} = \mathbf{lv} = \mathbf{0}$. The rate of deformation and spin are

$$\mathbf{d} = \frac{1}{2}\begin{bmatrix} 0 & \dot{\gamma} & 0 \\ \dot{\gamma} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{w} = \frac{1}{2}\begin{bmatrix} 0 & \dot{\gamma} & 0 \\ -\dot{\gamma} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and, from 2.5.14, $\dot{\gamma} = -\dot{\theta}_{12}$, the rate of change of the angle shown in Fig. 2.5.2.

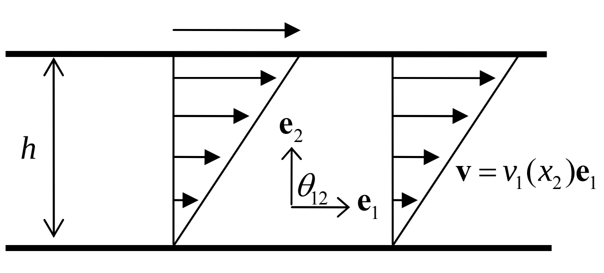


Figure 2.5.2: shear flow

The eigenvalues of $\mathbf{d}$ are $\lambda = 0, \pm\dot{\gamma}/2$ ($\det\mathbf{d} = 0$) and the principal invariants, Eqn. 1.11.17, are $\mathrm{I}_\mathbf{d} = 0$, $\mathrm{II}_\mathbf{d} = -\frac{1}{4}\dot{\gamma}^2$, $\mathrm{III}_\mathbf{d} = 0$. For $\lambda = +\dot{\gamma}/2$, the eigenvector is $\mathbf{n}_1 = [1 \quad 1 \quad 0]^T$ and for $\lambda = -\dot{\gamma}/2$, it is $\mathbf{n}_2 = [-1 \quad 1 \quad 0]^T$ (for $\lambda = 0$ it is $\mathbf{e}_3$). (The eigenvalues and eigenvectors of $\mathbf{w}$ are complex.) Relative to the basis of eigenvectors,

$$\mathbf{d} = \begin{bmatrix} \dot{\gamma}/2 & 0 & 0 \\ 0 & -\dot{\gamma}/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so at 45° there is an instantaneous pure rate of stretching/contraction of material.

■

2.5.4 Other Rates of Strain Tensors

From 2.2.9, 2.2.22,

$$\frac{1}{2}\frac{d}{dt}(d\mathbf{x}\cdot d\mathbf{x}) = d\mathbf{X}\frac{1}{2}\dot{\mathbf{C}}d\mathbf{X} = d\mathbf{X}\dot{\mathbf{E}}d\mathbf{X} \tag{2.5.16}$$

This can also be written in terms of spatial line elements:

$$d\mathbf{X}\dot{\mathbf{E}}d\mathbf{X} = d\mathbf{x}\left[\mathbf{F}^{-\mathrm{T}}\dot{\mathbf{E}}\mathbf{F}^{-1}\right]d\mathbf{x} \tag{2.5.17}$$

But from 2.5.9, these also equal $d\mathbf{x}\mathbf{d}d\mathbf{x}$, which leads to expressions for the material time derivatives of the right Cauchy-Green and Green-Lagrange strain tensors (also given here are expressions for the time derivatives of the left Cauchy-Green and Euler-Almansi tensors {▲Problem 3})

$$\boxed{\begin{aligned} &\dot{\mathbf{C}} = 2\mathbf{F}^{\mathrm{T}}\mathbf{d}\mathbf{F} \\ &\dot{\mathbf{E}} = \mathbf{F}^{\mathrm{T}}\mathbf{d}\mathbf{F} \\ &\dot{\mathbf{b}} = \mathbf{l}\mathbf{b} + \mathbf{b}\mathbf{l}^{\mathrm{T}} \\ &\dot{\mathbf{e}} = \mathbf{d} - \mathbf{l}^{\mathrm{T}}\mathbf{e} - \mathbf{e}\mathbf{l} \end{aligned}} \tag{2.5.18}$$

Note that

$$\int \dot{\mathbf{E}}\,dt = \int d\mathbf{E}$$

so that the integral of the rate of Green-Lagrange strain is path independent and, in particular, the integral of $\dot{\mathbf{E}}$ around any closed loop (so that the final configuration is the same as the initial configuration) is zero. However, in general, the integral of the rate of deformation,

$$\int \mathbf{d}dt$$

is not independent of the path – there is no universal function $\mathbf{h}$ such that $\mathbf{d} = d\mathbf{h}/dt$ with $\int \mathbf{d}dt = \int d\mathbf{h}$. Thus the integral $\int \mathbf{d}dt$ over a closed path may be non-zero, and hence the integral of the rate of deformation is not a good measure of the total strain.

The Hencky Strain

The Hencky strain is, Eqn. 2.2.37, $\mathbf{h} = \sum_{i=1}^{3} (\ln \lambda_i) \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i$, where $\mathbf{n}_i$ are the principal spatial axes. Thus, if the principal spatial axes do not change with time, $\dot{\mathbf{h}} = \sum_{i=1}^{3} (\dot{\lambda}_i / \lambda_i) \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i$. With the left stretch $\mathbf{v} = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i$, it follows that (and similarly for the corresponding material tensors), $\dot{\mathbf{H}} \equiv \dot{\overline{\ln \mathbf{U}}} = \dot{\mathbf{U}}\mathbf{U}^{-1}$, $\dot{\mathbf{h}} \equiv \dot{\overline{\ln \mathbf{v}}} = \dot{\mathbf{v}}\mathbf{v}^{-1}$.

For example, consider an extension in the coordinate directions, so $\mathbf{F} = \mathbf{U} = \mathbf{v} = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i = \sum_{i=1}^{3} \lambda_i \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i$. The motion and velocity are

$$x_i = \lambda_i X_i, \quad \dot{x}_i = \dot{\lambda}_i X_i = \frac{\dot{\lambda}_i}{\lambda_i} x_i \qquad (\text{no sum})$$

so $d_i = \dot{\lambda}_i / \lambda_i$ (no sum), and $\mathbf{d} = \dot{\mathbf{h}}$. Further, $\mathbf{h} = \int \mathbf{d} dt$. Note that, as mentioned above, this expression does not hold in general, but does in this case of uniform extension.

2.5.5 Material Derivatives of Line, Area and Volume Elements

The material derivative of a line element $d(d\mathbf{x})/dt$ has been derived (defined) through 2.4.8. For area and volume elements, it is necessary first to evaluate the material derivative of the Jacobian determinant *J*. From the chain rule, one has (see Eqns 1.15.11, 1.15.7)

$$\dot{J} = \frac{d}{dt}(J(\mathbf{F})) = \frac{\partial J}{\partial \mathbf{F}} : \dot{\mathbf{F}} = J\mathbf{F}^{-\mathrm{T}} : \dot{\mathbf{F}} \tag{2.5.19}$$

Hence {▲Problem 4}

$$\boxed{\begin{aligned} \dot{J} &= J \operatorname{tr}(\mathbf{l}) \\ &= J \operatorname{tr}(\operatorname{grad} \mathbf{v}) \\ &= J \operatorname{div} \mathbf{v} \end{aligned}} \tag{2.5.20}$$

Since $\mathbf{l} = \mathbf{d} + \mathbf{w}$ and $\operatorname{tr}\mathbf{w} = 0$, it also follows that $\dot{J} = J \operatorname{tr} \mathbf{d}$.

As mentioned earlier, an isochoric motion is one for which the volume is constant – thus any of the following statements characterise the necessary and sufficient conditions for an isochoric motion:

$$J = 1, \quad \dot{J} = 0, \quad \operatorname{div}\mathbf{v} = 0, \quad \operatorname{tr}\mathbf{d} = 0, \quad \mathbf{F}^{-\mathrm{T}} : \dot{\mathbf{F}} = 0 \tag{2.5.21}$$

Applying Nanson's formula 2.2.59, the material derivative of an area vector element is {▲Problem 6}

$$\boxed{\frac{d}{dt}(\hat{\mathbf{n}}ds) = \left(\operatorname{div}\mathbf{v} - \mathbf{l}^{\mathrm{T}}\right)\hat{\mathbf{n}}ds} \tag{2.5.22}$$

Finally, from 2.2.53, the material time derivative of a volume element is

$$\boxed{\frac{d}{dt}(dv) = \frac{d}{dt}(JdV) = \dot{J}dV = \operatorname{div}\mathbf{v}\,dv} \tag{2.5.23}$$

Example (Shear and Stretch)

Consider a sample of material undergoing the following motion, Fig. 2.4.3.

$$\begin{aligned} x_1 &= X_1 + k\lambda X_2 \\ x_2 &= \lambda X_2 \\ x_3 &= X_3 \end{aligned} \quad , \qquad \begin{aligned} X_1 &= x_1 - kx_2 \\ X_2 &= \frac{1}{\lambda}x_2 \\ X_3 &= x_3 \end{aligned}$$

with $\lambda = \lambda(t)$, $k = k(t)$.

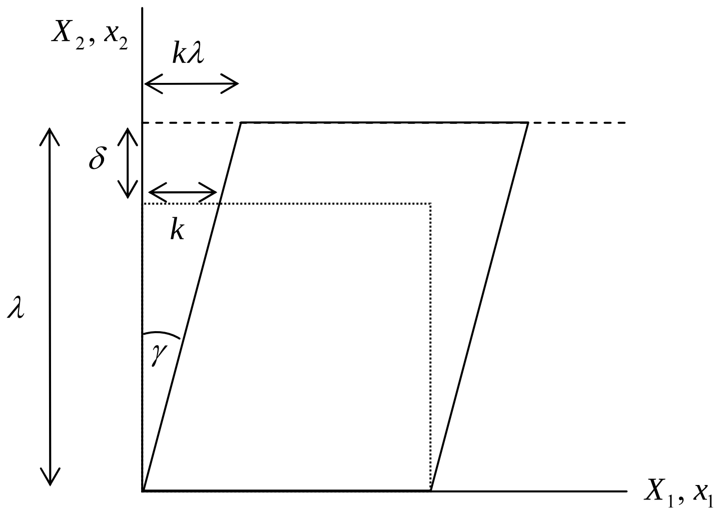


Figure 2.4.3: shear and stretch

The deformation gradient and material strain tensors are

$$\mathbf{F} = \begin{bmatrix} 1 & k\lambda & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & k\lambda & 0 \\ k\lambda & (1+k^2)\lambda^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 0 & \frac{1}{2}k\lambda & 0 \\ \frac{1}{2}k\lambda & \frac{1}{2}\left(\lambda^2(1+k^2)-1\right) & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

the Jacobian $J = \det\mathbf{F} = \lambda$, and the spatial strain tensors are

$$\mathbf{b} = \begin{bmatrix} 1+k^2\lambda^2 & k\lambda^2 & 0 \\ k\lambda^2 & \lambda^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{e} = \begin{bmatrix} 0 & \frac{1}{2}k & 0 \\ \frac{1}{2}k & \frac{1}{2}\dfrac{(1-k^2)\lambda^2 - 1}{\lambda^2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

This deformation can also be expressed as a stretch followed by a simple shear:

$$\mathbf{F} = \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The velocity is

$$\mathbf{V} = \frac{d\mathbf{x}}{dt} = \begin{bmatrix} (\dot{k}\lambda + k\dot{\lambda})X_2 \\ \dot{\lambda}X_2 \\ 0 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} (\dot{k} + k(\dot{\lambda}/\lambda))x_2 \\ (\dot{\lambda}/\lambda)x_2 \\ 0 \end{bmatrix}$$

The velocity gradient is

$$\mathbf{l} = \frac{d\mathbf{v}}{d\mathbf{x}} = \begin{bmatrix} 0 & \dot{k} + k(\dot{\lambda}/\lambda) & 0 \\ 0 & \dot{\lambda}/\lambda & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and the rate of deformation and spin are

$$\mathbf{d} = \begin{bmatrix} 0 & \frac{1}{2}\left[\dot{k} + k(\dot{\lambda}/\lambda)\right] & 0 \\ \frac{1}{2}\left[\dot{k} + k(\dot{\lambda}/\lambda)\right] & \dot{\lambda}/\lambda & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 0 & \frac{1}{2}\left[\dot{k} + k(\dot{\lambda}/\lambda)\right] & 0 \\ -\frac{1}{2}\left[\dot{k} + k(\dot{\lambda}/\lambda)\right] & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Also

$$\dot{\mathbf{C}} = 2\mathbf{F}^T\mathbf{dF} = \begin{bmatrix} 0 & \lambda\dot{k} + k\dot{\lambda} & 0 \\ \lambda\dot{k} + k\dot{\lambda} & 2\lambda\left(k\lambda\dot{k} + (k^2+1)\dot{\lambda}\right) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

As expected, from 2.5.20,

$$\dot{J} = J\mathrm{tr}(\mathbf{d}) = J(\dot{\lambda}/\lambda) = \dot{\lambda}$$

■

2.5.6 Problems

1. (a) Differentiate the relation $\mathbf{I} = \mathbf{F}\mathbf{F}^{-1}$ and use 2.5.4, $\dot{\mathbf{F}} = \mathbf{l}\mathbf{F}$, to derive 2.5.5b,

   $\dot{\overline{\mathbf{F}^{-1}}} = -\mathbf{F}^{-1}\mathbf{l}$.

   (b) Differentiate the relation $\mathbf{I} = \mathbf{F}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}}$ and use 2.5.4, $\dot{\mathbf{F}} = \mathbf{l}\mathbf{F}$, and 1.10.3e to derive 2.5.5c, $\dot{\overline{\mathbf{F}^{-\mathrm{T}}}} = -\mathbf{l}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}}$.

2. For the velocity field

   $$v_1 = x_1^2 x_2, \quad v_2 = 2x_2^2 x_3, \quad v_3 = 3x_1 x_2 x_3$$

   determine the rate of stretching per unit stretch at (2,0,1) in the direction of the unit vector

   $$(4\mathbf{e}_1 - 3\mathbf{e}_2)/5$$

   And in the direction of $\mathbf{e}_1$?

3. (a) Derive the relation 2.5.18a, $\dot{\mathbf{C}} = 2\mathbf{F}^{\mathrm{T}}\mathbf{d}\mathbf{F}$ directly from $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$

   (b) Use the definitions $\mathbf{b} = \mathbf{F}\mathbf{F}^{\mathrm{T}}$ and $\mathbf{e} = (\mathbf{I} - \mathbf{b}^{-1})/2$ to derive the relations 2.5.18c,d: $\dot{\mathbf{b}} = \mathbf{l}\mathbf{b} + \mathbf{b}\mathbf{l}^{\mathrm{T}}$, $\quad \dot{\mathbf{e}} = \mathbf{d} - \mathbf{l}^{\mathrm{T}}\mathbf{e} - \mathbf{e}\mathbf{l}$

4. Use 2.5.4, 2.5.19, 1.10.3h, 1.10.6, to derive 2.5.20.

5. For the motion $x_1 = 3X_1 t - t^2,\ x_2 = X_1 + X_2 t,\ x_3 = tX_3$, verify that $\dot{\mathbf{F}} = \mathbf{l}\mathbf{F}$. What is the ratio of the volume element currently occupying $(1,1,1)$ to its volume in the undeformed configuration? And what is the rate of change of this volume element, per unit current volume?

6. Use Nanson's formula 2.2.59, the product rule of differentiation, and 2.5.20, 2.5.5c, to derive the material time derivative of a vector area element, 2.5.22 (note that $\hat{\mathbf{N}}$, a unit normal in the undeformed configuration, is constant).

2.6 Deformation Rates: Further Topics

2.6.1 Relationship between l, d, w and the rate of change of R and U

Consider the polar decomposition $\mathbf{F} = \mathbf{RU}$. Since $\mathbf{R}$ is orthogonal, $\mathbf{RR}^{\mathrm{T}} = \mathbf{I}$, and a differentiation of this equation leads to

$$\mathbf{\Omega}_{\mathbf{R}} \equiv \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}} = -\mathbf{R}\dot{\mathbf{R}}^{\mathrm{T}} \tag{2.6.1}$$

with $\mathbf{\Omega}_{\mathbf{R}}$ skew-symmetric (see Eqn. 1.14.2). Using this relation, the expression $\mathbf{l} = \dot{\mathbf{F}}\mathbf{F}^{-1}$, and the definitions of **d** and **w**, Eqn. 2.5.7, one finds that {▲Problem 1}

$$\begin{aligned}
\mathbf{l} &= \mathbf{R}\dot{\mathbf{U}}\mathbf{U}^{-1}\mathbf{R}^{\mathrm{T}} + \mathbf{\Omega}_{\mathbf{R}} \\
\mathbf{w} &= \frac{1}{2}\mathbf{R}\left(\dot{\mathbf{U}}\mathbf{U}^{-1} - \mathbf{U}^{-1}\dot{\mathbf{U}}\right)\mathbf{R}^{\mathrm{T}} + \mathbf{\Omega}_{\mathbf{R}} \\
&= \mathbf{R}\,\mathrm{skew}\left[\dot{\mathbf{U}}\mathbf{U}^{-1}\right]\mathbf{R}^{\mathrm{T}} + \mathbf{\Omega}_{\mathbf{R}} \\
\mathbf{d} &= \frac{1}{2}\mathbf{R}\left(\dot{\mathbf{U}}\mathbf{U}^{-1} + \mathbf{U}^{-1}\dot{\mathbf{U}}\right)\mathbf{R}^{\mathrm{T}} \\
&= \mathbf{R}\,\mathrm{sym}\left[\dot{\mathbf{U}}\mathbf{U}^{-1}\right]\mathbf{R}^{\mathrm{T}}
\end{aligned} \tag{2.6.2}$$

Note that $\mathbf{\Omega}_{\mathbf{R}}$ being skew-symmetric is consistent with **w** being skew-symmetric, and that both **w** and **d** involve **R**, and the rate of change of **U**.

When the motion is a rigid body rotation, then $\dot{\mathbf{U}} = \mathbf{0}$, and

$$\mathbf{w} = \mathbf{\Omega}_{\mathbf{R}} = \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}} \tag{2.6.3}$$

2.6.2 Deformation Rate Tensors and the Principal Material and Spatial Bases

The rate of change of the stretch tensor in terms of the principal material base vectors is

$$\dot{\mathbf{U}} = \sum_{i=1}^{3}\left\{\dot{\lambda}_i\hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i + \lambda_i\dot{\hat{\mathbf{N}}}_i \otimes \hat{\mathbf{N}}_i + \lambda_i\hat{\mathbf{N}}_i \otimes \dot{\hat{\mathbf{N}}}_i\right\} \tag{2.6.4}$$

Consider the case when the principal material axes stay constant, as can happen in some simple deformations. In that case, $\dot{\mathbf{U}}$ and $\mathbf{U}^{-1}$ are coaxial (see §1.11.5):

$$\dot{\mathbf{U}} = \sum_{i=1}^{3}\dot{\lambda}_i\hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i \quad \text{and} \quad \mathbf{U}^{-1} = \sum_{i=1}^{3}\frac{1}{\lambda_i}\hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i \tag{2.6.5}$$

with $\dot{\mathbf{U}}\mathbf{U}^{-1} = \mathbf{U}^{-1}\dot{\mathbf{U}}$ and, as expected, from 2.5.25b, $\mathbf{w} = \mathbf{\Omega}_\mathbf{R} = \dot{\mathbf{R}}\mathbf{R}^\mathrm{T}$, that is, any spin is due to rigid body rotation.

Similarly, from 2.2.37, and differentiating $\hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i = \mathbf{I}$,

$$\dot{\mathbf{E}} = \sum_{i=1}^{3} \left\{ \lambda_i \dot{\lambda}_i \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i + \tfrac{1}{2}\lambda_i^2 \dot{\hat{\mathbf{N}}}_i \otimes \hat{\mathbf{N}}_i + \tfrac{1}{2}\lambda_i^2 \hat{\mathbf{N}}_i \otimes \dot{\hat{\mathbf{N}}}_i \right\}. \tag{2.6.6}$$

Also, differentiating $\hat{\mathbf{N}}_i \cdot \hat{\mathbf{N}}_j = \delta_{ij}$ leads to $\dot{\hat{\mathbf{N}}}_i \cdot \hat{\mathbf{N}}_j = -\hat{\mathbf{N}}_i \cdot \dot{\hat{\mathbf{N}}}_j$ and so the expression

$$\dot{\hat{\mathbf{N}}}_i = \sum_{m=1}^{3} W_{im} \hat{\mathbf{N}}_m \tag{2.6.7}$$

is valid provided W_{ij} are the components of a skew-symmetric tensor, $W_{ij} = -W_{ji}$. This leads to an alternative expression for the Green-Lagrange tensor:

$$\dot{\mathbf{E}} = \sum_{i=1}^{3} \lambda_i \dot{\lambda}_i \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i + \sum_{\substack{m,n=1 \\ m\neq n}}^{3} \tfrac{1}{2} W_{mn} \left(\lambda_m^2 - \lambda_n^2\right) \hat{\mathbf{N}}_m \otimes \hat{\mathbf{N}}_n \tag{2.6.8}$$

Similarly, from 2.2.37, the left Cauchy-Green tensor can be expressed in terms of the principal spatial base vectors:

$$\mathbf{b} = \sum_{i=1}^{3} \lambda_i^2 \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i, \quad \dot{\mathbf{b}} = \sum_{i=1}^{3} \left\{ 2\lambda_i \dot{\lambda}_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i + \lambda_i^2 \dot{\hat{\mathbf{n}}}_i \otimes \hat{\mathbf{n}}_i + \lambda_i^2 \hat{\mathbf{n}}_i \otimes \dot{\hat{\mathbf{n}}}_i \right\} \tag{2.6.9}$$

Then, from inspection of 2.5.18c, $\dot{\mathbf{b}} = \mathbf{lb} + \mathbf{bl}^\mathrm{T}$, the velocity gradient can be expressed as {▲Problem 2}

$$\mathbf{l} = \sum_{i=1}^{3} \left\{ \frac{\dot{\lambda}_i}{\lambda_i} \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i + \dot{\hat{\mathbf{n}}}_i \otimes \hat{\mathbf{n}}_i \right\} = \sum_{i=1}^{3} \left\{ \frac{\dot{\lambda}_i}{\lambda_i} \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i - \hat{\mathbf{n}}_i \otimes \dot{\hat{\mathbf{n}}}_i \right\} \tag{2.6.7}$$

2.6.3 Rates of Change and the Relative Deformation

Just as the material time derivative of the deformation gradient is defined as

$$\dot{\mathbf{F}} = \frac{\partial}{\partial t} \mathbf{F}(\mathbf{X}, t) = \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{x}}{\partial \mathbf{X}} \right)$$

one can define the material time derivative of the relative deformation gradient, *cf.* §2.3.2, the rate of change *relative to the current configuration*:

$$\dot{\mathbf{F}}_t(\mathbf{x}, t) = \frac{\partial}{\partial \tau} \mathbf{F}_t(\mathbf{x}, \tau)\Big|_{\tau = t} \tag{2.6.8}$$

From 2.3.8, $\mathbf{F}_t(\mathbf{x},\tau) = \mathbf{F}(\mathbf{X},\tau)\mathbf{F}(\mathbf{X},t)^{-1}$, so taking the derivative with respect to τ (t is now fixed) and setting $\tau = t$ gives

$$\dot{\mathbf{F}}_t(\mathbf{x},t) = \dot{\mathbf{F}}(\mathbf{X},t)\mathbf{F}(\mathbf{X},t)^{-1}$$

Then, from 2.5.4,

$$\mathbf{l} = \dot{\mathbf{F}}_t(\mathbf{x},t) \tag{2.6.9}$$

as expected – the velocity gradient is the rate of change of deformation relative to the current configuration. Further, using the polar decomposition,

$$\mathbf{F}_t(\mathbf{x},\tau) = \mathbf{R}_t(\mathbf{x},\tau)\mathbf{U}_t(\mathbf{x},\tau)$$

Differentiating with respect to τ and setting $\tau = t$ then gives

$$\dot{\mathbf{F}}_t(\mathbf{x},t) = \mathbf{R}_t(\mathbf{x},t)\dot{\mathbf{U}}_t(\mathbf{x},t) + \dot{\mathbf{R}}_t(\mathbf{x},t)\mathbf{U}_t(\mathbf{x},t)$$

Relative to the current configuration, $\mathbf{R}_t(\mathbf{x},t) = \mathbf{U}_t(\mathbf{x},t) = \mathbf{I}$, so, from 2.4.34,

$$\mathbf{l} = \dot{\mathbf{U}}_t(\mathbf{x},t) + \dot{\mathbf{R}}_t(\mathbf{x},t) \tag{2.6.10}$$

With **U** symmetric and **R** skew-symmetric, $\dot{\mathbf{U}}_t(\mathbf{x},t)$, $\dot{\mathbf{R}}_t(\mathbf{x},t)$ are, respectively, symmetric and skew-symmetric, and it follows that

$$\begin{aligned} \mathbf{d} &= \dot{\mathbf{U}}_t(\mathbf{x},t) \\ \mathbf{w} &= \dot{\mathbf{R}}_t(\mathbf{x},t) \end{aligned} \tag{2.6.11}$$

again, as expected – the rate of deformation is the instantaneous rate of stretching and the spin is the instantaneous rate of rotation.

The Corotational Derivative

The **corotational derivative** of a vector **a** is $\overset{\circ}{\mathbf{a}} \equiv \dot{\mathbf{a}} - \mathbf{w}\mathbf{a}$. Formally, it is defined through

$$\begin{aligned} \overset{\circ}{\mathbf{a}} &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\mathbf{a}(t+\Delta t) - \mathbf{R}_t(t+\Delta t)\mathbf{a}(t)\right\} \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\mathbf{a}(t+\Delta t) - \left[\mathbf{R}_t(t) + \Delta t\dot{\mathbf{R}}_t(t) + \cdots\right]\mathbf{a}(t)\right\} \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\mathbf{a}(t+\Delta t) - \left[\mathbf{I} + \Delta t\mathbf{w}(t) + \cdots\right]\mathbf{a}(t)\right\} \\ &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\mathbf{a}(t+\Delta t) - \mathbf{a}(t)\right\} - \mathbf{w}(t)\mathbf{a}(t) \\ &= \dot{\mathbf{a}} - \mathbf{w}\mathbf{a} \end{aligned} \tag{2.6.12}$$

The definition shows that the corotational derivative involves taking a vector **a** in the current configuration and rotating it with the rigid body rotation part of the motion, Fig. 2.6.1. It is this new, rotated, vector which is compared with the vector $\mathbf{a}(t+\Delta t)$, which has undergone rotation and stretch.

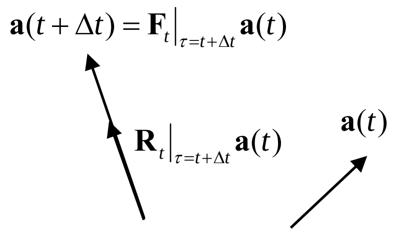


Figure 2.6.1: rotation and stretch of a vector

2.6.4 Rivlin-Ericksen Tensors

The n-th **Rivlin-Ericksen tensor** is defined as

$$\mathbf{A}_n(t) = \left.\frac{d^n}{d\tau^n}\mathbf{C}_t(\tau)\right|_{\tau=t}, \qquad n = 0, 1, 2, \cdots \tag{2.6.13}$$

where $\mathbf{C}_t(\tau)$ is the relative right Cauchy-Green strain. Since $\mathbf{C}_t(\tau)\big|_{\tau=t} = \mathbf{I}$, $\mathbf{A}_0 = \mathbf{I}$. To evaluate the next Rivlin-Ericksen tensor, one needs the derivatives of the relative deformation gradient; from 2.5.4, 2.3.8,

$$\frac{d}{d\tau}\mathbf{F}_t(\tau) = \frac{d}{d\tau}\left[\mathbf{F}(\tau)\mathbf{F}(t)^{-1}\right] = \mathbf{l}(\tau)\mathbf{F}(\tau)\mathbf{F}(t)^{-1} = \mathbf{l}(\tau)\mathbf{F}_t(\tau) \tag{2.6.14}$$

Then, with 2.5.5a, $d\left(\mathbf{F}_t(\tau)^{\mathrm{T}}\right)/d\tau = \mathbf{F}_t(\tau)^{\mathrm{T}}\mathbf{l}(\tau)^{\mathrm{T}}$, and

$$\begin{aligned}\mathbf{A}_1(t) &= \left[\mathbf{F}_t(\tau)^{\mathrm{T}}\left(\mathbf{l}(\tau)+\mathbf{l}(\tau)^{\mathrm{T}}\right)\mathbf{F}_t(\tau)\right]_{\tau=t} \\ &= \left(\mathbf{l}(t)+\mathbf{l}(t)^{\mathrm{T}}\right) \\ &= 2\mathbf{d}\end{aligned}$$

Thus the tensor $\mathbf{A}_1$ gives a measure of the rate of stretching of material line elements (see Eqn. 2.5.10). Similarly, higher Rivlin-Ericksen tensors give a measure of higher order stretch rates, $\ddot{\lambda}, \dddot{\lambda}$, and so on.

2.6.5 The Directional Derivative and the Material Time Derivative

The directional derivative of a function $\mathbf{T}(t)$ in the direction of an increment in t is, by definition (see, for example, Eqn. 1.15.27),

$$\partial_t \mathbf{T}[\Delta t] = \mathbf{T}(t + \Delta t) - \mathbf{T}(t) \tag{2.6.15}$$

or

$$\partial_t \mathbf{T}[\Delta t] = \frac{d\mathbf{T}}{dt} \Delta t \tag{2.6.16}$$

Setting $\Delta t = 1$, and using the chain rule 1.15.28,

$$\begin{aligned} \dot{\mathbf{T}} &= \partial_t \mathbf{T}[1] \\ &= \partial_{\mathbf{x}} \mathbf{T}\left[\partial_t \mathbf{x}[1]\right] \\ &= \partial_{\mathbf{x}} \mathbf{T}\left[\mathbf{v}\right] \end{aligned} \tag{2.6.17}$$

The material time derivative is thus equivalent to the directional derivative in the direction of the velocity vector.

2.6.6 Problems

1. Derive the relations 2.6.2.
2. Use 2.6.9 to verify 2.5.18, $\dot{\mathbf{b}} = \mathbf{l}\mathbf{b} + \mathbf{b}\mathbf{l}^{\mathrm{T}}$.

2.7 Small Strain Theory

When the deformation is small, from 2.2.43-4,

$$\begin{aligned}\mathbf{F} &= \mathbf{I} + \mathrm{Grad}\mathbf{U} \\ &= \mathbf{I} + (\mathrm{grad}\mathbf{u})\mathbf{F} \\ &\approx \mathbf{I} + \mathrm{grad}\mathbf{u}\end{aligned} \tag{2.7.1}$$

neglecting the product of $\mathrm{grad}\mathbf{u}$ with $\mathrm{Grad}\mathbf{U}$, since these are small quantities. Thus one can take $\mathrm{Grad}\mathbf{U} = \mathrm{grad}\mathbf{u}$ and there is no distinction to be made between the undeformed and deformed configurations. The deformation gradient is of the form $\mathbf{F} = \mathbf{I} + \boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is small.

2.7.1 Decomposition of Strain

Any second order tensor can be decomposed into its symmetric and antisymmetric part according to 1.10.28, so that

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial \mathbf{x}} &= \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^{\mathrm{T}}\right) + \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}} - \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^{\mathrm{T}}\right) = \boldsymbol{\varepsilon} + \boldsymbol{\Omega} \\ \frac{\partial u_i}{\partial x_j} &= \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right) \quad = \varepsilon_{ij} + \Omega_{ij}\end{aligned} \tag{2.7.2}$$

where $\boldsymbol{\varepsilon}$ is the small strain tensor 2.2.48 and $\boldsymbol{\Omega}$, the anti-symmetric part of the displacement gradient, is the **small rotation tensor**, so that **F** can be written as

$$\boxed{\mathbf{F} = \mathbf{I} + \boldsymbol{\varepsilon} + \boldsymbol{\Omega}} \quad \textbf{Small Strain Decomposition of the Deformation Gradient} \tag{2.7.3}$$

It follows that (for the calculation of **e**, one can use the relation $(\mathbf{I} + \boldsymbol{\delta})^{-1} \approx \mathbf{I} - \boldsymbol{\delta}$ for small $\boldsymbol{\delta}$)

$$\begin{aligned}\mathbf{C} &= \mathbf{b} = \mathbf{I} + 2\boldsymbol{\varepsilon} \\ \mathbf{E} &= \mathbf{e} = \boldsymbol{\varepsilon}\end{aligned} \tag{2.7.4}$$

Rotation

Since $\boldsymbol{\Omega}$ is antisymmetric, it can be written in terms of an axial vector $\boldsymbol{\omega}$, *cf.* §1.10.11, so that for any vector **a**,

$$\boldsymbol{\Omega}\mathbf{a} = \boldsymbol{\omega} \times \mathbf{a}, \qquad \boldsymbol{\omega} = -\Omega_{23}\mathbf{e}_1 + \Omega_{13}\mathbf{e}_1 - \Omega_{12}\mathbf{e}_3 \tag{2.7.5}$$

The relative displacement can now be written as

$$\begin{aligned} d\mathbf{u} &= (\mathrm{grad}\,\mathbf{u})d\mathbf{X} \\ &= \boldsymbol{\varepsilon} d\mathbf{X} + \boldsymbol{\omega} \times d\mathbf{X} \end{aligned} \tag{2.7.6}$$

The component of relative displacement given by $\boldsymbol{\omega} \times d\mathbf{X}$ is perpendicular to $d\mathbf{X}$, and so represents a pure rotation of the material line element, Fig. 2.7.1.

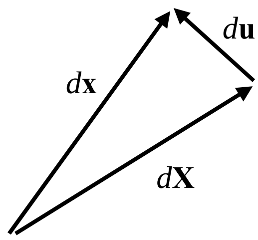


Figure 2.7.1: a pure rotation

Principal Strains

Since $\boldsymbol{\varepsilon}$ is symmetric, it must have three mutually orthogonal eigenvectors, the **principal axes of strain**, and three corresponding real eigenvalues, the **principal strains,** e_1, e_2, e_3), which can be positive or negative, *cf.* §1.11. The effect of $\boldsymbol{\varepsilon}$ is therefore to deform an elemental unit sphere into an elemental ellipsoid, whose axes are the principal axes, and whose lengths are $1+e_1, 1+e_2, 1+e_3$. Material fibres in these principal directions are stretched only, in which case the deformation is called a **pure deformation**; fibres in other directions will be stretched and rotated.

The term $\boldsymbol{\varepsilon} d\mathbf{X}$ in 2.7.6 therefore corresponds to a pure stretch along the principal axes. The total deformation is the sum of a pure deformation, represented by $\boldsymbol{\varepsilon}$, and a rigid body rotation, represented by $\boldsymbol{\Omega}$. This result is similar to that obtained for the exact finite strain theory, but here the decomposition is *additive* rather than *multiplicative*. Indeed, here the corresponding small strain stretch and rotation tensors are $\mathbf{U} = \mathbf{I} + \boldsymbol{\varepsilon}$ and $\mathbf{R} = \mathbf{I} + \boldsymbol{\Omega}$, so that

$$\mathbf{F} = \mathbf{RU} = \mathbf{I} + \boldsymbol{\varepsilon} + \boldsymbol{\Omega} \tag{2.7.7}$$

Example

Consider the simple shear (*c.f.* Eqn. 2.2.40)

$$x_1 = X_1 + kX_2, \quad x_2 = X_2, \quad x_3 = X_3$$

where k is small. The displacement vector is $\mathbf{u} = kx_2\mathbf{e}_1$ so that

$$\mathrm{grad}\,\mathbf{u} = \begin{bmatrix} 0 & k & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The deformation can be written as the additive decomposition

$$d\mathbf{u} = \boldsymbol{\varepsilon} d\mathbf{X} + \boldsymbol{\Omega} d\mathbf{X} \qquad \text{or} \qquad d\mathbf{u} = \boldsymbol{\varepsilon} d\mathbf{X} + \boldsymbol{\omega} \times d\mathbf{X}$$

with

$$\boldsymbol{\varepsilon} = \begin{bmatrix} 0 & k/2 & 0 \\ k/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} 0 & k/2 & 0 \\ -k/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and $\boldsymbol{\omega} = -(k/2)\mathbf{e}_3$. For the rotation component, one can write

$$\mathbf{R} = \mathbf{I} + \boldsymbol{\Omega} = \begin{bmatrix} 1 & k/2 & 0 \\ -k/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which, since for small θ, $\cos\theta \approx 1, \sin\theta \approx \theta$, can be seen to be a rotation through an angle $\theta = -k/2$ (a clockwise rotation).

The principal values of $\boldsymbol{\varepsilon}$ are $\pm k/2, 0$ with corresponding principal directions $\mathbf{n}_1 = (1/\sqrt{2})\mathbf{e}_1 + (1/\sqrt{2})\mathbf{e}_2$, $\mathbf{n}_2 = -(1/\sqrt{2})\mathbf{e}_1 + (1/\sqrt{2})\mathbf{e}_2$ and $\mathbf{n}_3 = \mathbf{e}_3$.

Thus the simple shear with small displacements consists of a rotation through an angle $k/2$ superimposed upon a pure shear with angle $k/2$, Fig. 2.6.2.

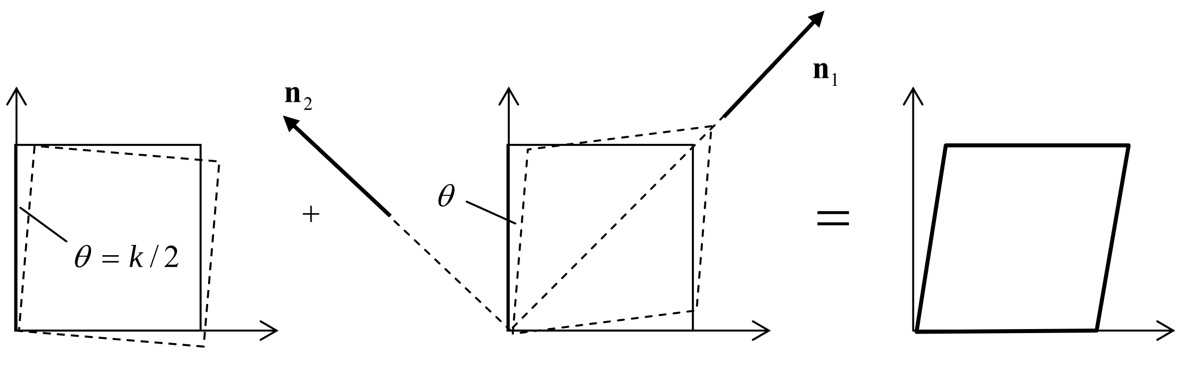


Figure 2.6.2: simple shear

■

2.7.2 Rotations and Small Strain

Consider now a pure rotation about the X_3 axis (within the exact finite strain theory), $d\mathbf{x} = \mathbf{R}d\mathbf{X}$, with

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.7.8}$$

This rotation does not change the length of line elements $d\mathbf{X}$. According to the small strain theory, however,

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \cos\theta - 1 & 0 & 0 \\ 0 & \cos\theta - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} 0 & -\sin\theta & 0 \\ \sin\theta & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which does predict line element length changes, but which can be neglected if θ is small. For example, if the rotation is of the order 10^{-2} rad, then $\varepsilon_{11} = \varepsilon_{22} = 10^{-4}$. However, if the rotation is large, the errors will be appreciable; in that case, rigid body rotation introduces geometrical non-linearities which must be dealt with using the finite deformation theory.

Thus the small strain theory is restricted to not only the case of small displacement gradients, but also small rigid body rotations.

2.7.3 Volume Change

An elemental cube with edges of unit length in the directions of the principal axes deforms into a cube with edges of lengths $1+e_1, 1+e_2, 1+e_3$, so the unit change in volume of the cube is

$$\frac{dv - dV}{dV} = (1+e_1)(1+e_2)(1+e_3) - 1 = e_1 + e_2 + e_3 + O(2) \tag{2.7.9}$$

Since second order quantities have already been neglected in introducing the small strain tensor, they must be neglected here. Hence the increase in volume per unit volume, called the **dilatation** (or **dilation**) is

$$\boxed{\frac{\delta V}{V} = e_1 + e_2 + e_3 = e_{ii} = \mathrm{tr}\boldsymbol{\varepsilon} = \mathrm{div}\mathbf{u}} \quad \textbf{Dilatation} \tag{2.7.10}$$

Since any elemental volume can be constructed out of an infinite number of such elemental cubes, this result holds for any elemental volume irrespective of shape.

2.7.4 Rate of Deformation, Strain Rate and Spin Tensors

Take now the expressions 2.4.7 for the rate of deformation and spin tensors. Replacing $\mathbf{v}$ in these expressions by $\dot{\mathbf{u}}$, one has

$$\mathbf{d} = \frac{1}{2}\left(\mathbf{l} + \mathbf{l}^{\mathrm{T}}\right), \qquad d_{ij} = \frac{1}{2}\left(\frac{\partial \dot{u}_i}{\partial x_j} + \frac{\partial \dot{u}_j}{\partial x_i}\right)$$
$$\mathbf{w} = \frac{1}{2}\left(\mathbf{l} - \mathbf{l}^{\mathrm{T}}\right), \qquad w_{ij} = \frac{1}{2}\left(\frac{\partial \dot{u}_i}{\partial x_j} - \frac{\partial \dot{u}_j}{\partial x_i}\right) \tag{2.7.11}$$

For small strains, one can take the time derivative outside (by considering the x_i to be material coordinates independent of time):

$$d_{ij} = \frac{d}{dt}\left\{\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right\}$$
$$w_{ij} = \frac{d}{dt}\left\{\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right)\right\} \tag{2.7.12}$$

The rate of deformation in this context is seen to be the **rate of strain**, $\mathbf{d} = \dot{\boldsymbol{\varepsilon}}$, and the spin is seen to be the **rate of rotation**, $\mathbf{w} = \dot{\boldsymbol{\Omega}}$.

The instantaneous motion of a material particle can hence be regarded as the sum of three effects:

(i) a translation given by $\dot{\mathbf{u}}$ (so in the time interval Δt the particle has been displaced by $\dot{\mathbf{u}}\Delta t$)
(ii) a pure deformation given by $\dot{\boldsymbol{\varepsilon}}$
(iii) a rigid body rotation given by $\dot{\boldsymbol{\Omega}}$

2.7.5 Compatibility Conditions

Suppose that the strains ε_{ij} in a body are known. If the displacements are to be determined, then the strain-displacement partial differential equations

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{2.7.13}$$

need to be integrated. However, there are six independent strain components but only three displacement components. This implies that the strains are not independent but are related in some way. The relations between the strains are called **compatibility conditions**, and it can be shown that they are given by

$$\varepsilon_{ij,km} + \varepsilon_{km,ij} - \varepsilon_{ik,jm} - \varepsilon_{jm,ik} = 0 \tag{2.7.14}$$

These are 81 equations, but only six of them are distinct, and these six equations are necessary and sufficient to evaluate the displacement field.

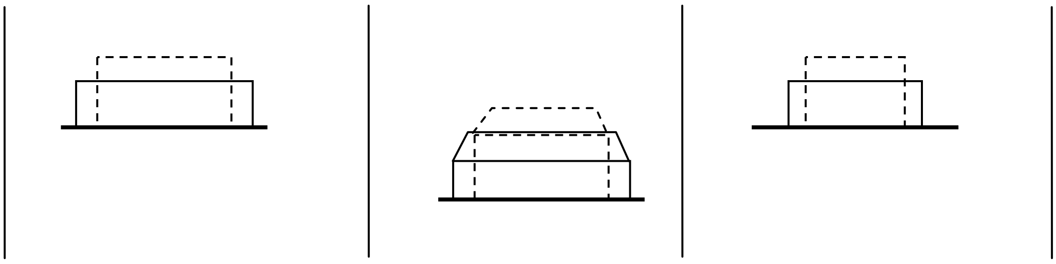
2.8 Objectivity and Objective Tensors

2.8.1 Dependence on Observer

Consider a rectangular block of material resting on a circular table. A person stands and observes the material deform, Fig. 2.8.1a. The dashed lines indicate the undeformed material whereas the solid line indicates the current state. A second observer is standing just behind the first, but on a step ladder – this observer sees the material as in 2.8.1b. A third observer is standing around the table, 45° from the first, and sees the material as in Fig. 2.8.1c.

The deformation can be described by each observer using concepts like displacement, velocity, strain and so on.. However, it is clear that the three observers will in general record different values for these measures, since their perspectives differ.

The goal in what follows is to determine which of the kinematical tensors are in fact *independent* of observer. Since the laws of physics describing the response of a deforming material must be independent of any observer, it is these particular tensors which will be more readily used in expressions to describe material response.



(a) (b) (c)

Figure 2.8.1: a deforming material as seen by different observers

Note that Fig. 2.8.1 can be interpreted in another, equivalent, way. One can imagine *one* static observer, but this time with the material moved into three different positions. This viewpoint will be returned to in the next section.

2.8.2 Change of Reference Frame

Consider two **frames of reference**, the first consisting of the origin $\mathbf{o}$ and the basis $\{\mathbf{e}_i\}$, the second consisting of the origin $\mathbf{o}^*$ and the basis $\{\mathbf{e}_i^*\}$, Fig. 2.8.2. A point $\mathbf{x}$ in space is then identified as having position vector $\mathbf{x} = x_i \mathbf{e}_i$ in the first frame and position vector $\mathbf{x}^* = x_i^* \mathbf{e}_i^*$ in the second frame.

When the origins $\mathbf{o}$ and $\mathbf{o}^*$ coincide, $\mathbf{x} = \mathbf{x}^*$ and the vector components x_i and x_i^* are related through Eqn. 1.5.3, $x_i = Q_{ij} x_j^*$, or $\mathbf{x} = x_i \mathbf{e}_i = Q_{ij} x_j^* \mathbf{e}_i$, where $[\mathbf{Q}]$ is the transformation matrix 1.5.4, $Q_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j^*$. Alternatively, one has Eqn. 1.5.5, $x_i^* = Q_{ji} x_j$, or $\mathbf{x}^* = x_i^* \mathbf{e}_i^* = Q_{ji} x_j \mathbf{e}_i^*$.

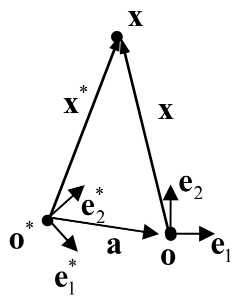


Figure 2.8.2: two frames of reference

With the shift in origin $\mathbf{a} = \mathbf{o} - \mathbf{o}^*$, one has

$$\mathbf{x}^* = x_i^* \mathbf{e}_i^* = Q_{ji} x_j \mathbf{e}_i^* + a_i^* \mathbf{e}_i^* \tag{2.8.1}$$

where $\mathbf{a} = a_i^* \mathbf{e}_i^*$. Alternatively,

$$\mathbf{x} = x_i \mathbf{e}_i = Q_{ij} x_j^* \mathbf{e}_i - a_i \mathbf{e}_i \tag{2.8.2}$$

where $\mathbf{a} = a_i \mathbf{e}_i$, with $a_i^* = Q_{ji} a_j$.

Formulae 2.8.1-2 relate the coordinates of the position vector to a point in space as observed from one frame of reference to the coordinates of the position vector to the *same* point as observed from a different frame of reference.

Finally, consider the position vector **x**, which is defined relative to the frame $(\mathbf{o}, \mathbf{e}_i)$. To an observer in the frame $(\mathbf{o}^*, \mathbf{e}_i^*)$, the *same* position vector would appear as $(\mathbf{x})^*$, Fig. 2.8.3. Rotating this vector $(\mathbf{x})^*$ through $\mathbf{Q}^{\mathrm{T}}$ (the tensor which rotates the basis $\{\mathbf{e}_i^*\}$ into the basis $\{\mathbf{e}_i\}$) and adding the vector **a** then produces $\mathbf{x}^*$:

$$\mathbf{x}^* = \mathbf{Q}^{\mathrm{T}} (\mathbf{x})^* + \mathbf{a} \tag{2.8.3}$$

This relation will be discussed further below.

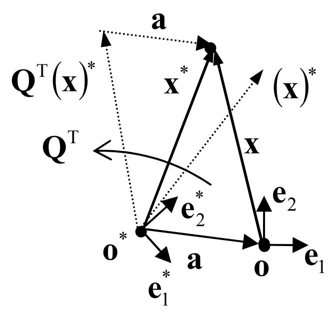


Figure 2.8.3: Relation between vectors in Eqn. 2.8.3

2.8.3 Change of Observer

The change of frame encompassed by Eqns. 2.8.1-2 is more precisely called a **passive change of frame**, and merely involves a transformation between vector components. One would say that there is one observer but that this observer is using two frames of reference. Here follows a different concept, an **active change of frame**, also called a **change in observer**, in which there are two observers, each with their own frame of reference.

An **observer** is someone who can measure relative positions in space (with a ruler) and instants of time (with a clock). An **event** in the physical world (for example a material particle) is perceived by an observer as occurring at a particular point in space and at a particular time. One can regard an observer O to be a map of an event E in the physical world to a point $\mathbf{x}$ in point space (*cf.* §1.2.5) and a real number t. A *single* event E is recorded as the pair $(\mathbf{x}, t)$ by an observer O and, in general, by a *different* pair $(\mathbf{x}^*, t^*)$ by a second observer O^*, Fig. 2.8.4.

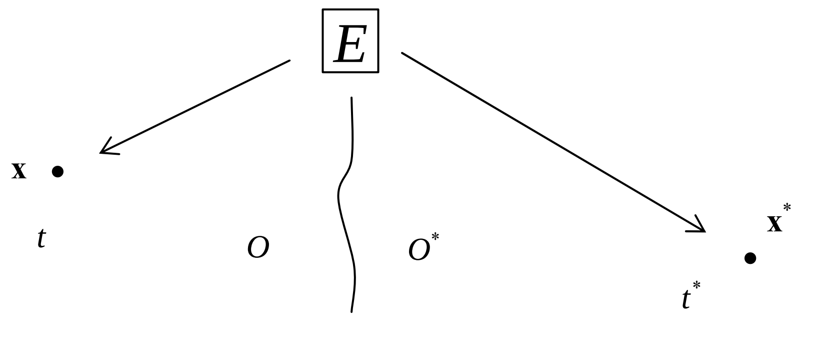


Figure 2.8.4: recordings by two observers of the same event

Let the two observers record three points corresponding to three events, Fig. 2.8.5. These points define vectors in space, as the difference between the points (*cf.* §1.2.5). It is assumed that both observers "see" the same Euclidean geometry, that is, if one observer sees an ellipse, then the other observer will see the same ellipse, but perhaps positioned differently in space. To ensure that this is so, observed vectors must be related through some orthogonal tensor $\mathbf{Q}$, for example,

$$\mathbf{x}^* - \mathbf{x}_0^* = \mathbf{Q}(\mathbf{x} - \mathbf{x}_0) \tag{2.8.4}$$

since this transformation will automatically preserve distances between points, and angles between vectors (see §1.10.7), for example,

$$\left(\mathbf{x}_1^* - \mathbf{x}_0^*\right)\cdot\left(\mathbf{x}^* - \mathbf{x}_0^*\right) = \mathbf{Q}(\mathbf{x}_1 - \mathbf{x}_0)\cdot\mathbf{Q}(\mathbf{x} - \mathbf{x}_0) = (\mathbf{x}_1 - \mathbf{x}_0)\cdot(\mathbf{x} - \mathbf{x}_0) \qquad (2.8.5)$$

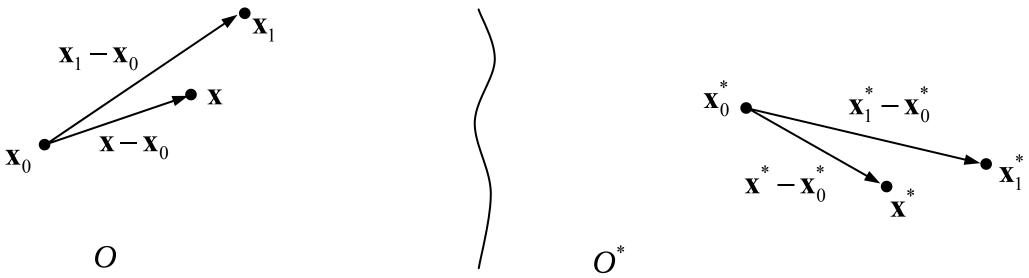


Figure 2.8.5: recordings of two observers of three separate events

Although all orthogonal tensors $\mathbf{Q}$ do indeed preserve length and angles, it is taken that the $\mathbf{Q}$ in 2.8.4-5 is proper orthogonal, i.e. a rotation tensor (*cf.* §1.10.8), so that orientation is also preserved. Further, it is assumed that $\mathbf{Q} = \mathbf{Q}(t)$, which expresses the fact that the observers can move relative to each other over time.

Observers must also agree on time intervals between events. Let an observer O record a certain event at time t and a second observer O^* record the same event as occurring at time t^*. Then the times must be related through

$$\boxed{t^* = t + \alpha} \qquad \textbf{Observer Time Transformation} \qquad (2.8.6)$$

where α is a *constant*. If now the observers record a second event as occurring at t_1 and t_1^* say, one has $t_1^* - t^* = t_1 - t$ as required.

The observer transformation 2.8.4 involves the vectors $\mathbf{x} - \mathbf{x}_0$ and $\mathbf{x}^* - \mathbf{x}_0^*$ and as such does not require the notion of origin or coordinate system; it is an abstract symbolic notation for an observer transformation. However, an origin $\mathbf{o}$ for O and $\mathbf{o}^*$ for O^* can be introduced and then the points $\mathbf{x}_0, \mathbf{x}, \mathbf{x}_0^*, \mathbf{x}^*$ can be regarded as *position vectors* in space, Fig. 2.8.6.

The transformation 2.8.4 can now be expressed in the oft-used format

$$\boxed{\mathbf{x}^* = \mathbf{c}(t) + \mathbf{Q}(t)\mathbf{x}} \qquad \textbf{Observer (Spatial) Transformation} \qquad (2.8.7)$$

where

$$\mathbf{c}(t) = \mathbf{x}_0^* - \mathbf{Q}(t)\mathbf{x}_0 \qquad (2.8.8)$$

The transformation 2.8.7 is called a **Euclidean transformation**, since it preserves the Euclidean geometry.

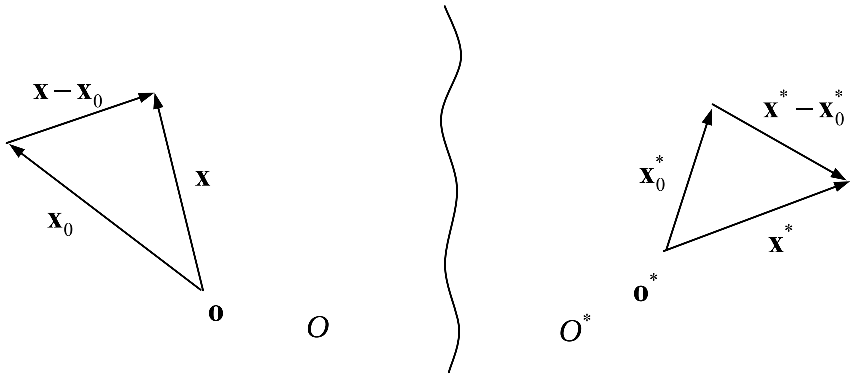


Figure 2.8.6: position vectors for two observers of the same events

Coordinate Systems

Each observer can introduce any Cartesian coordinate system, with basis vectors $\{\mathbf{e}_i\}$ and $\{\mathbf{e}_i^*\}$ say. They can then resolve the position vectors into vector components. These basis vectors can be oriented with respect to each other in any way, that is, they will be related through $\mathbf{e}_i^* = \mathbf{R}\mathbf{e}_i$, where $\mathbf{R}$ is any rotation tensor. Indeed, each observer can change their basis, effecting a coordinate transformation. No attempt to introduce specific coordinate systems will be made here since they are completely unnecessary to the notion of observer transformation and would only greatly confuse the issue.

Relationship to Passive Change of Frame

Recall the passive change of frame encompassed in Eqns. 2.8.1-2. If one substitutes the *actual* $\mathbf{x}$ for $(\mathbf{x})^*$ in Eqn. 2.8.3, one has:

$$\mathbf{x}^* = \mathbf{Q}^{\mathrm{T}}\mathbf{x} + \mathbf{a} \tag{2.8.9}$$

This is clearly an observer transformation, relating the position vector as seen by one observer to the position vector as seen by a second observer, through an orthogonal tensor and a vector, as in Eqn. 2.8.7. In the passive change of frame, Q_{ij} are the components of the orthogonal tensor $\mathbf{Q} = \mathbf{e}_i^* \otimes \mathbf{e}_i$, Eqn. 1.10.25, which maps the bases onto each other: $\mathbf{e}_i^* = \mathbf{Q}\mathbf{e}_i$. Thus the transformation 2.8.1-2 can be defined uniquely by the pair $\mathbf{Q}$ and $\mathbf{a}$. In that sense, the passive change of frame does indeed define an active change of frame, i.e. a change of observer, through Eqn. 2.8.9. However, the concept of observer discussed above is the preferred way of defining an observer transformation.

2.8.4 Objective Vectors and Tensors

The observer transformation 2.8.7 encapsulates the different viewpoints observers have of the physical world. They will see the same objects, but in general they will see these objects oriented differently and located at different positions. The goal now is to see which of the kinematical tensors are independent of these different viewpoints. As a first step, next is introduced the concept of an **objective tensor**.

Suppose that different observers are examining a deforming material. In order to describe the material, the observers take measurements. This will involve measurements of *spatial* objects associated with the current configuration, for example the velocity or spin. It will also involve *material* objects associated with the reference configuration, for example line elements in that configuration. It will also involve *two-point* tensors such as the rotation or deformation gradient, which are associated with both the current and reference configurations.

It is assumed that all observers observe the reference configuration to be the same, that is, they record the same set of points for the material particles in the reference configuration[1]. The observers then move relative to each other and their measurements of objects associated with the current configuration will in general differ. One would expect (want) different observers to make the same measurement of material objects despite this relative movement; thus one says that material vectors and tensors are **objective (material) vectors** and **objective (material) tensors** if they remain unchanged under the observer transformation 2.8.6-7.

A spatial vector $\mathbf{u}$ on the other hand is said to be an **objective (spatial) vector** if it satisfies the observer transformation (see 2.8.4):[2]

$$\boxed{\mathbf{u}^* = \mathbf{Qu}} \qquad \textbf{Objectivity Requirement for a Spatial Vector} \qquad (2.8.10)$$

for all rotation tensors $\mathbf{Q}$. An **objective (spatial) tensor** is defined to be one which transforms an objective vector into an objective vector. Consider a tensor observed as $\mathbf{T}$ and $\mathbf{T}^*$ by two different observers. Take an objective vector which is observed as $\mathbf{v}$ and $\mathbf{v}^*$, and let $\mathbf{u} = \mathbf{Tv}$ and $\mathbf{u}^* = \mathbf{T}^*\mathbf{v}^*$. Then, for $\mathbf{u}$ to be objective,

$$\mathbf{u}^* = \mathbf{Qu} = \mathbf{QTv} = \mathbf{QTQ}^{\mathrm{T}}\mathbf{v}^* \qquad (2.8.11)$$

and so the tensor is objective provided

$$\boxed{\mathbf{T}^* = \mathbf{QTQ}^{\mathrm{T}}} \qquad \textbf{Objectivity Requirement for a Spatial Tensor} \qquad (2.8.12)$$

Various identities can be derived; for example, for objective vectors $\mathbf{a}$ and $\mathbf{b}$, and objective tensors $\mathbf{A}$ and $\mathbf{B}$, {▲Problem 1}

---

[1] this does not affect the generality of what follows; the notion of objective tensor is independent of the chosen reference configuration
[2] the time transformation 2.8.6 is trivial and does not affect the relations to be derived

$$\begin{aligned}
(\mathbf{a}+\mathbf{b})^* &= \mathbf{a}^* + \mathbf{b}^* \\
(\mathbf{a}\otimes\mathbf{b})^* &= \mathbf{a}^* \otimes \mathbf{b}^* \\
(\mathbf{a}\cdot\mathbf{b})^* &= \mathbf{a}^* \cdot \mathbf{b}^* \\
(\mathbf{A}\mathbf{b})^* &= \mathbf{A}^*\mathbf{b}^* \\
(\mathbf{A}\mathbf{B})^* &= \mathbf{A}^*\mathbf{B}^* \\
(\mathbf{A}^{-1})^* &= (\mathbf{A}^*)^{-1} \\
(\mathbf{A}\mathbf{B})^* &= \mathbf{A}^*\mathbf{B}^* \\
(\mathbf{A}:\mathbf{B})^* &= \mathbf{A}^* : \mathbf{B}^*
\end{aligned} \tag{2.8.13}$$

For a scalar,

$$\boxed{\phi^* = \phi} \qquad \textbf{Objectivity Requirement for a Scalar} \tag{2.8.14}$$

In other words, an **objective scalar** is one which has the same value to all observers.

Finally, consider a two-point tensor. Such a tensor is said to be objective if it maps an objective material vector into an objective spatial vector. Consider then a two-point tensor observed as $\mathbf{T}$ and $\mathbf{T}^*$. Take an objective material vector which is observed as $\mathbf{v}$ and $\mathbf{v}^*$, and let $\mathbf{u} = \mathbf{T}\mathbf{v}$ and $\mathbf{u}^* = \mathbf{T}^*\mathbf{v}^*$. A material vector is objective if it is unaffected by an observer transformation, so

$$\mathbf{u}^* = \mathbf{Q}\mathbf{u} = \mathbf{Q}\mathbf{T}\mathbf{v} = \mathbf{Q}\mathbf{T}\mathbf{v}^* \tag{2.8.15}$$

and so the tensor is objective provided

$$\boxed{\mathbf{T}^* = \mathbf{Q}\mathbf{T}} \qquad \textbf{Objectivity Requirement for a Two-point Tensor} \tag{2.8.16}$$

Thus the objectivity requirement for a two-point tensor is the same as that for a spatial vector.

2.8.5 Objective Kinematics

Next are examined the various kinematic vectors and tensors introduced in the earlier sections, and their objectivity status is determined.

The motion is observed by one observer as $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$ and by a second observer as $\mathbf{x}^* = \boldsymbol{\chi}(\mathbf{X}, t^*)$. The observer transformation gives

$$\boldsymbol{\chi}^*(\mathbf{X}, t^*) = \mathbf{Q}(t)\boldsymbol{\chi}(\mathbf{X}, t) + \mathbf{c}(t), \qquad t^* = t + \alpha \tag{2.8.17}$$

and so the motion is not an objective vector, i.e. $\boldsymbol{\chi}^* \neq \mathbf{Q}\boldsymbol{\chi}$.

The Velocity and Acceleration

Differentiating 2.8.17 (and using the notation $\dot{\mathbf{x}}$ instead of $\dot{\boldsymbol{\chi}}(\mathbf{X}, t)$ for brevity), the velocity under the observer transformation is

$$\dot{\mathbf{x}}^* = \dot{\mathbf{Q}}\mathbf{x} + \mathbf{Q}\dot{\mathbf{x}} + \dot{\mathbf{c}} \tag{2.8.18}$$

which does not comply with the objectivity requirement for spatial vectors, 2.8.10. In other words, different observers will measure different magnitudes for the velocity. The velocity expression can be put in a form similar to that of elementary mechanics (the "non-objective" terms are on the right),

$$\dot{\mathbf{x}}^* - \mathbf{Q}\dot{\mathbf{x}} = \boldsymbol{\Omega}_{\mathbf{Q}}\left(\mathbf{x}^* - \mathbf{c}\right) + \dot{\mathbf{c}} \tag{2.8.19}$$

where

$$\boldsymbol{\Omega}_{\mathbf{Q}} = \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} \tag{2.8.20}$$

is skew-symmetric (see Eqn. 1.14.2); this tensor represents the rigid body angular velocity between the observers (see Eqn. 2.6.1). Note that the velocity *is* objective provided $\dot{\mathbf{Q}} = \mathbf{0}$, $\dot{\mathbf{c}} = \mathbf{o}$, for which $\mathbf{x}^* = \mathbf{Q}_0\mathbf{x} + \mathbf{c}_0$, which is called a **time-independent rigid transformation**.

Similarly, for the acceleration, it can be shown that

$$\ddot{\mathbf{x}}^* - \mathbf{Q}\ddot{\mathbf{x}} = \dot{\boldsymbol{\Omega}}_{\mathbf{Q}}\left(\mathbf{x}^* - \mathbf{c}\right) - \boldsymbol{\Omega}_{\mathbf{Q}}^2\left(\mathbf{x}^* - \mathbf{c}\right) + 2\boldsymbol{\Omega}_{\mathbf{Q}}\left(\dot{\mathbf{x}} - \dot{\mathbf{c}}\right) + \ddot{\mathbf{c}} \tag{2.8.21}$$

The first three terms on the right-hand side are called the **Euler acceleration**, the **centrifugal acceleration** and the **Coriolis acceleration** respectively. The acceleration is objective provided $\dot{\mathbf{c}}$ and $\mathbf{Q}$ are constant, for which $\mathbf{x}^* = \mathbf{Q}_0\mathbf{x} + \mathbf{c}(t)$ with $\ddot{\mathbf{c}} = \mathbf{o}$, which is called a **Galilean transformation** – where the two configurations are related by a rigid rotation and a translational motion with constant velocity.

The Deformation Gradient

Consider the motion $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t)$. As mentioned, observers observe the reference configuration to be the same: $\mathbf{X}^* = \mathbf{X}$. The deformation is then observed as $d\mathbf{x} = \mathbf{F}d\mathbf{X}$ and $d\mathbf{x}^* = \mathbf{F}^* d\mathbf{X}$, so that

$$d\mathbf{x}^* = \mathbf{Q}d\mathbf{x} = \mathbf{Q}\mathbf{F}d\mathbf{X} = \mathbf{Q}\mathbf{F}d\mathbf{X} \tag{2.8.22}$$

and

$$\mathbf{F}^* = \mathbf{Q}\mathbf{F} \tag{2.8.23}$$

and so, according to 2.8.16, the deformation gradient is objective.

The Cauchy-Green Strain Tensors

For the right and left Cauchy-Green tensors,

$$\begin{aligned} \mathbf{C}^* &= \mathbf{F}^{*\mathrm{T}}\mathbf{F}^* = \mathbf{F}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\mathbf{F} = \mathbf{C} \\ \mathbf{b}^* &= \mathbf{F}^*\mathbf{F}^{*\mathrm{T}} = \mathbf{Q}\mathbf{F}\mathbf{F}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}\mathbf{b}\mathbf{Q}^{\mathrm{T}} \end{aligned} \tag{2.8.24}$$

Thus the material tensor **C** and the spatial tensor **b** are objective[3].

The Jacobian Determinant

For the Jacobian determinant, using 1.10.16a,

$$J^* = \det \mathbf{F}^* = \det(\mathbf{QF}) = \det \mathbf{Q} \det \mathbf{F} = \det \mathbf{F} = J \tag{2.8.25}$$

and[4] so is objective according to 2.8.14.

The Rotation and Stretch Tensors

The polar decomposition is $\mathbf{F} = \mathbf{RU}$, where **R** is the orthogonal rotation tensor and **U** is the right stretch tensor. Then $\mathbf{F}^* = \mathbf{QF} = \mathbf{QRU} \equiv \mathbf{R}^*\mathbf{U}^*$. Since $\mathbf{QR}$ is orthogonal, the expression $\mathbf{QRU} = \mathbf{R}^*\mathbf{U}^*$ is valid provided

$$\mathbf{R}^* = \mathbf{QR}, \qquad \mathbf{U}^* = \mathbf{U} \tag{2.8.26}$$

Thus the two-point tensor **R** and the material tensor **U** are objective.

The Velocity Gradient

Allowing **Q** to be a function of time, for the velocity gradient, using 2.5.4, 1.9.18c,

$$\mathbf{l}^* = \dot{\overline{\mathbf{F}^*}}(\mathbf{F}^*)^{-1} = (\mathbf{Q}\dot{\mathbf{F}} + \dot{\mathbf{Q}}\mathbf{F})\mathbf{F}^{-1}\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}\mathbf{l}\mathbf{Q}^{\mathrm{T}} + \mathbf{\Omega}_{\mathbf{Q}} \tag{2.8.27}$$

where $\mathbf{\Omega}_{\mathbf{Q}}$ is the angular velocity tensor 2.8.20. On the other hand, with $\mathbf{l} = \mathbf{d} + \mathbf{w}$, and separating out the symmetric and skew-symmetric parts,

$$\mathbf{d}^* = \mathbf{Q}\mathbf{d}\mathbf{Q}^{\mathrm{T}}, \qquad \mathbf{w}^* = \mathbf{Q}\mathbf{w}\mathbf{Q}^{\mathrm{T}} + \mathbf{\Omega}_{\mathbf{Q}} \tag{2.8.28}$$

Thus the velocity gradient is not objective. This is not surprising given that the velocity is not objective. However, significantly, the rate of deformation, a measure of the rate of stretching of material, *is* objective.

---

[3] Some authors define a second order tensor to be objective only if 2.8.12 is satisfied, regardless of whether it is spatial, two-point or material; with this definition, **F** and **C** would be defined as non-objective
[4] Note that **Q** must be a rotation tensor, not just an orthogonal tensor, here

The Spatial Gradient

Consider the spatial gradient of an *objective vector* **t**:

$$\operatorname{grad}\mathbf{t} = \frac{\partial \mathbf{t}}{\partial \mathbf{x}}, \qquad (\operatorname{grad}\mathbf{t})^* = \frac{\partial \mathbf{t}^*}{\partial \mathbf{x}^*} \tag{2.8.29}$$

Since $\mathbf{t}^* = \mathbf{Qt}$, the chain rule gives

$$\frac{\partial \mathbf{t}^*}{\partial \mathbf{x}} = \frac{\partial \mathbf{t}^*}{\partial \mathbf{x}^*}\frac{\partial \mathbf{x}^*}{\partial \mathbf{x}} \equiv \frac{\partial (\mathbf{Qt})}{\partial \mathbf{x}} = \mathbf{Q}\frac{\partial \mathbf{t}}{\partial \mathbf{x}} \tag{2.8.30}$$

It follows that

$$(\operatorname{grad}\mathbf{t})^* = \mathbf{Q}\frac{\partial \mathbf{t}}{\partial \mathbf{x}}\mathbf{Q}^{\mathrm{T}} \tag{2.8.31}$$

Thus the spatial gradient is objective. In general, it can be shown that the spatial gradient of a tensor field of order *n* is objective, for example the gradient of a scalar ϕ, {▲Problem 2} $\operatorname{grad}\phi$. Further, for a vector **v**, {▲Problem 3} $\operatorname{div}\mathbf{v}$ is objective.

Objective Rates

Consider an objective vector field **u**. The material derivative $\dot{\mathbf{u}}$ is not objective. However, the co-rotational derivative, Eqn. 2.6.12, $\overset{\circ}{\mathbf{u}} = \dot{\mathbf{u}} - \mathbf{wu}$ *is* objective. To show this, contract 2.8.28b, $\mathbf{w}^* = \mathbf{QwQ}^{\mathrm{T}} + \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}$, to the right with $\mathbf{Q}$ to get an expression for $\dot{\mathbf{Q}}$:

$$\dot{\mathbf{Q}} = \mathbf{w}^*\mathbf{Q} - \mathbf{Qw} \tag{2.8.32}$$

and then

$$\mathbf{u}^* = \mathbf{Qu} \quad \to \quad \dot{\overline{\mathbf{u}^*}} = \dot{\mathbf{Q}}\mathbf{u} + \mathbf{Q}\dot{\mathbf{u}} = \mathbf{w}^*\mathbf{Qu} + \mathbf{Q}(\dot{\mathbf{u}} - \mathbf{wu}) = \mathbf{w}^*\mathbf{Qu} + \mathbf{Q}\overset{\circ}{\mathbf{u}} \tag{2.8.33}$$

Then $\dot{\overline{\mathbf{u}^*}} - \mathbf{w}^*\mathbf{u}^* = \mathbf{Q}\overset{\circ}{\mathbf{u}}$, or $(\overset{\circ}{\mathbf{u}})^* = \mathbf{Q}\overset{\circ}{\mathbf{u}}$, so that the co-rotational derivative of a vector is an objective vector.

Rates of spatial tensors can also be modified in order to construct objective rates. For example, consider an objective spatial tensor **T**, so $\mathbf{T}^* = \mathbf{QTQ}^{\mathrm{T}}$. Then

$$\dot{\overline{\mathbf{T}^*}} = \mathbf{Q}\dot{\mathbf{T}}\mathbf{Q}^{\mathrm{T}} + \dot{\mathbf{Q}}\mathbf{TQ}^{\mathrm{T}} + \mathbf{QT}\dot{\mathbf{Q}}^{\mathrm{T}} \tag{2.8.34}$$

which is clearly not objective. However, this can be re-arranged using 2.8.32 into

$$\overline{\mathbf{T}^*}^{\cdot} - \mathbf{w}^*\mathbf{T}^* + \mathbf{T}^*\mathbf{w}^* = \mathbf{Q}\left(\dot{\mathbf{T}} - \mathbf{w}\mathbf{T} + \mathbf{T}\mathbf{w}\right)\mathbf{Q}^{\mathrm{T}} \tag{2.8.35}$$

and so the quantity

$$\dot{\mathbf{T}} - \mathbf{w}\mathbf{T} + \mathbf{T}\mathbf{w} \tag{2.8.36}$$

is an objective rate, called the **Jaumann rate**. Other objective rates of tensors can be constructed in a similar fashion, for example the **Cotter-Rivlin rate**, defined by {▲Problem 4}

$$\dot{\mathbf{T}} + \mathbf{l}^{\mathrm{T}}\mathbf{T} + \mathbf{T}\mathbf{l} \tag{2.8.37}$$

Summary of Objective Kinematic Objects

Table 2.8.1 summarises the objectivity of some important kinematic objects:

	objective	definition	Type	Transformation
Jacobian determinant	✓		Scalar	$J^* = J$
Deformation gradient	✓		2-point	$\mathbf{F}^* = \mathbf{Q}\mathbf{F}$
Rotation	✓	$\mathbf{R} = \mathbf{F}\mathbf{U}^{-1} = \mathbf{v}^{-1}\mathbf{F}$	2-point	$\mathbf{R}^* = \mathbf{Q}\mathbf{R}$
Right Cauchy-Green strain	✓	$\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$	Material	$\mathbf{C}^* = \mathbf{C}$
Green-Lagrange strain	✓	$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$	Material	$\mathbf{E}^* = \mathbf{E}$
Rate of Green-Lagrange strain	✓		Material	$\overline{\mathbf{E}^*}^{\cdot} = \dot{\mathbf{E}}$
Right Stretch	✓	$\mathbf{U} = \sqrt{\mathbf{C}}$	Material	$\mathbf{U}^* = \mathbf{U}$
Left Cauchy-Green strain	✓	$\mathbf{b} = \mathbf{F}\mathbf{F}^{\mathrm{T}}$	Spatial	$\mathbf{b}^* = \mathbf{Q}\mathbf{b}\mathbf{Q}^{\mathrm{T}}$
Euler-Almansi strain	✓	$\mathbf{e} = \frac{1}{2}(\mathbf{I} - \mathbf{b}^{-1})$	Spatial	$\mathbf{e}^* = \mathbf{Q}\mathbf{e}\mathbf{Q}^{\mathrm{T}}$
Left Stretch	✓	$\mathbf{v} = \sqrt{\mathbf{b}}$	Spatial	$\mathbf{v}^* = \mathbf{Q}\mathbf{v}\mathbf{Q}^{\mathrm{T}}$
Spatial Velocity Gradient	×	$\mathbf{l} = \operatorname{grad}\mathbf{v}$	Spatial	$\mathbf{l}^* = \mathbf{Q}\mathbf{l}\mathbf{Q}^{\mathrm{T}} + \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}$
Rate of Deformation	✓	$\mathbf{d} = \frac{1}{2}(\mathbf{l} + \mathbf{l}^{\mathrm{T}})$	Spatial	$\mathbf{d}^* = \mathbf{Q}\mathbf{d}\mathbf{Q}^{\mathrm{T}}$
Spin	×	$\mathbf{w} = \frac{1}{2}(\mathbf{l} - \mathbf{l}^{\mathrm{T}})$	Spatial	$\mathbf{w}^* = \mathbf{Q}\mathbf{w}\mathbf{Q}^{\mathrm{T}} + \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}$

Table 2.8.1: Objective kinematic objects

2.8.6 Objective Functions

In a similar way, functions are defined to be objective as follows:

- A scalar-valued function ϕ of, for example, a tensor **A**, is objective if it transforms in the same way as an objective scalar,

$$\phi^*(\mathbf{A}) = \phi(\mathbf{A}) \tag{2.8.38}$$

- A (spatial) vector-valued function **a** of a tensor **A** is objective if it transforms in the same way as an objective vector

$$\mathbf{v}^*(\mathbf{A}) = \mathbf{Q}\mathbf{v}(\mathbf{A}) \tag{2.8.39}$$

- A (spatial) tensor-valued function **f** of a tensor **A** is objective if it transforms according to

$$\mathbf{f}^*(\mathbf{A}) = \mathbf{Q}\mathbf{f}(\mathbf{A})\mathbf{Q}^{\mathrm{T}} \tag{2.8.40}$$

Objective functions of the Deformation Gradient

Consider an objective scalar-valued function ϕ of the deformation gradient **F**, $\phi(\mathbf{F})$. The function is objective if $\phi^* = \phi(\mathbf{F})$. But also,

$$\phi^* = \phi(\mathbf{F}^*) = \phi(\mathbf{QF}) \tag{2.8.41}$$

Using the polar decomposition theorem, $\phi(\mathbf{RU}) = \phi(\mathbf{QRU})$. Choosing the particular rigid-body rotation $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$ then leads to

$$\phi(\mathbf{RU}) = \phi(\mathbf{U}) \tag{2.8.42}$$

which leads to the **reduced form**

$$\phi(\mathbf{F}) = \phi(\mathbf{U}) \tag{2.8.43}$$

Thus for the scalar function ϕ to be objective, it must be independent of the rotational part of **F**, and depends only on the stretching part; it cannot be a function of the nine independent components of the deformation gradient, but only of the six independent components of the right stretch tensor.

Consider next an objective (spatial) tensor-valued function **f** of the deformation gradient **F**, $\mathbf{f}(\mathbf{F})$. According to the definition of objectivity of a second order tensor, 2.8.12:

$$\mathbf{f}^* = \mathbf{Q}\mathbf{f}(\mathbf{F})\mathbf{Q}^{\mathrm{T}} \tag{2.8.44}$$

But also,

$$\mathbf{f}^* = \mathbf{f}(\mathbf{F}^*) = \mathbf{f}(\mathbf{QF}) \tag{2.8.45}$$

Again, using the polar decomposition theorem and choosing the particular rigid-body rotation $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$ leads to

$$\mathbf{f}(\mathbf{U}) = \mathbf{R}^{\mathrm{T}}\mathbf{f}(\mathbf{RU})\mathbf{R} \tag{2.8.46}$$

which leads to the reduced form

$$\mathbf{f}(\mathbf{F}) = \mathbf{R}\mathbf{f}(\mathbf{U})\mathbf{R}^{\mathrm{T}} \tag{2.8.47}$$

Thus for **f** to be objective, its dependence on **F** must be through an arbitrary function of **U** together with a more explicit dependence on **R**, the rotation tensor

Example

Consider the tensor function $\mathbf{f}(\mathbf{F}) = \alpha\left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right)^2$. Then

$$\mathbf{f}(\mathbf{Q}\mathbf{F}) = \alpha\left[(\mathbf{Q}\mathbf{F})(\mathbf{Q}\mathbf{F})^{\mathrm{T}}\right]^2 = \mathbf{Q}\alpha\left[\mathbf{F}\mathbf{F}^{\mathrm{T}}\right]^2\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}\mathbf{f}(\mathbf{F})\mathbf{Q}^{\mathrm{T}}$$

and so the objectivity requirement is satisfied. According to the above, then, one can evaluate $\mathbf{f}(\mathbf{U}) = \mathbf{R}^{\mathrm{T}}\mathbf{f}(\mathbf{R}\mathbf{U})\mathbf{R} = \alpha\left(\mathbf{U}\mathbf{U}^{\mathrm{T}}\right)^2$, and the reduced form is

$$\mathbf{f} = \mathbf{R}\alpha\left(\mathbf{U}\mathbf{U}^{\mathrm{T}}\right)^2\mathbf{R}^{\mathrm{T}} = \alpha\mathbf{R}\mathbf{U}^4\mathbf{R}^{\mathrm{T}}$$

Also, since $\mathbf{C} = \mathbf{U}^2$ and $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$, alternative reduced forms are

$$\mathbf{f} = \mathbf{R}\mathbf{f}_2(\mathbf{C})\mathbf{R}^{\mathrm{T}}, \quad \mathbf{f} = \mathbf{R}\mathbf{f}_3(\mathbf{E})\mathbf{R}^{\mathrm{T}}$$

■

Finally, consider a *spatial* tensor function **f** of a *material* tensor **T**. Then

$$\mathbf{f}^*(\mathbf{T}) = \mathbf{Q}\mathbf{f}(\mathbf{T})\mathbf{Q}^{\mathrm{T}}, \qquad \mathbf{f}^*(\mathbf{T}) = \mathbf{f}(\mathbf{T}^*) = \mathbf{f}(\mathbf{T}) \tag{2.8.48}$$

It follows that

$$\mathbf{f} = \mathbf{Q}\mathbf{f}\mathbf{Q}^{\mathrm{T}} \tag{2.8.49}$$

This is true only in the special case $\mathbf{Q} = \mathbf{I}$ and so is not true in general. It follows that the function **f** is not objective.

2.8.7 Problems

1. Derive the relations 2.8.13
2. Show that the spatial gradient of a scalar ϕ is objective.
3. Show that the divergence of a spatial vector **v** is objective. [Hint: use the definition 1.11.9 and identity 1.9.10e]
4. Verify that the Rivlin-Cotter rate of a tensor **T**, $\dot{\mathbf{T}} + \mathbf{l}^{\mathrm{T}}\mathbf{T} + \mathbf{T}\mathbf{l}$, is objective.

2.9 Rigid Body Rotations of Configurations

In this section are discussed rigid body rotations to the current and reference configurations.

2.9.1 A Rigid Body Rotation of the Current Configuration

As mentioned in §2.8.1, the circumstance of two observers, moving relative to each other and examining a fixed configuration (the current configuration) is equivalent to one observer taking measurements of two different configurations, moving relative to each other[1]. The objectivity requirements of the various kinematic objects discussed in the previous section can thus also be examined by considering rigid body rotations and translations of the current configuration.

Any rigid body rotation and translation of the current configuration can be expressed in the form

$$\mathbf{x}^*(\mathbf{X},t) = \mathbf{Q}(t)\mathbf{x}(\mathbf{X},t) + \mathbf{c}(t) \tag{2.9.1}$$

where $\mathbf{Q}$ is a rotation tensor. This is illustrated in Fig. 2.9.5. The current configuration is denoted by S and the rotated configuration by S^*.

Just as $d\mathbf{x} = \mathbf{F}d\mathbf{X}$, the deformation gradient for the configuration S^* relative to the reference configuration S_0 is defined through $d\mathbf{x}^* = \mathbf{F}^* d\mathbf{X}$. From 2.9.1, as in §2.8.5 (see Eqn. 2.8.23), and similarly for the right and left Cauchy-Green tensors,

$$\begin{aligned} \mathbf{F}^* &= \mathbf{QF} \\ \mathbf{C}^* &= \mathbf{F}^{*\mathrm{T}}\mathbf{F}^* = \mathbf{C} \\ \mathbf{b}^* &= \mathbf{F}^*\mathbf{F}^{*\mathrm{T}} = \mathbf{QbQ}^{\mathrm{T}} \end{aligned} \tag{2.9.2}$$

Thus in the deformations $\mathbf{F}: S_0 \to S$ and $\mathbf{F}^*: S_0 \to S^*$, the right Cauchy Green tensors, $\mathbf{C}$ and $\mathbf{C}^*$, are the same, but the left Cauchy Green tensors are different, and related through $\mathbf{b}^* = \mathbf{QbQ}^{\mathrm{T}}$.

All the other results obtained in the last section in the context of observer transformations, for example for the Jacobian, stretch tensors, etc., hold also for the case of rotations to the current configuration.

[1] Although equivalent, there is a difference: in one, there are two observers who record one event (a material particle say) as at two different points, in the other there is one observer who records two different events (the place where the one material particle is in two different configurations)

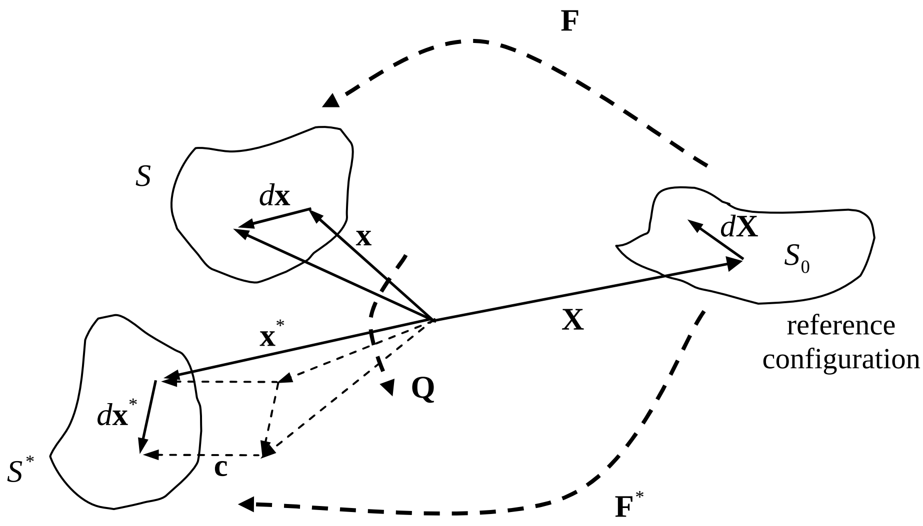


Figure 2.9.1: a rigid body rotation and translation of the current configuration

2.9.2 A Rigid Body Rotation of the Reference Configuration

Consider now a rigid-body rotation to the *reference* configuration. Such rotations play an important role in the notion of material symmetry (see Chapter 5).

The reference configuration is denoted by S_0 and the rotated/translated configuration by $S^\lozenge$, Fig. 2.9.2. The deformation gradient for the current configuration S relative to $S^\lozenge$ is defined through $d\mathbf{x} = \mathbf{F}^\lozenge d\mathbf{X}^\lozenge = \mathbf{F}^\lozenge \mathbf{Q} d\mathbf{X}$. But $d\mathbf{x} = \mathbf{F} d\mathbf{X}$ and so (and similarly for the right and left Cauchy-Green tensors)

$$\begin{aligned} \mathbf{F}^\lozenge &= \mathbf{F}\mathbf{Q}^{\mathrm{T}} \\ \mathbf{C}^\lozenge &= \mathbf{F}^{\lozenge\mathrm{T}}\mathbf{F}^\lozenge = \mathbf{Q}\mathbf{C}\mathbf{Q}^{\mathrm{T}} \\ \mathbf{b}^\lozenge &= \mathbf{F}^\lozenge\mathbf{F}^{\lozenge\mathrm{T}} = \mathbf{b} \end{aligned} \tag{2.9.3}$$

Thus the change to the right (left) Cauchy-Green strain tensor under a rotation to the reference configuration is the same as the change to the left (right) Cauchy-Green strain tensor under a rotation of the current configuration.

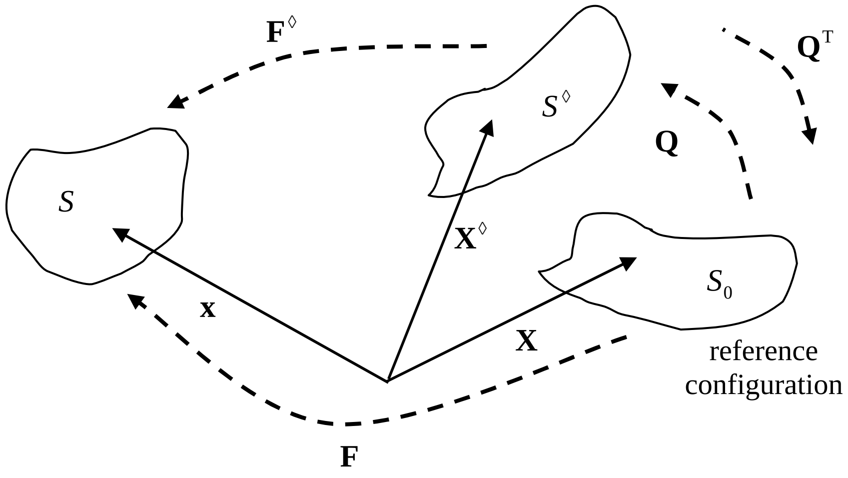


Figure 2.9.2: a rigid body rotation of the reference configuration

2.10 Convected Coordinates

An introduction to curvilinear coordinate was given in section 1.16, which serves as an introduction to this section. As mentioned there, the formulation of almost all mechanics problems, and their numerical implementation and solution, can be achieved using a description of the problem in terms of Cartesian coordinates. However, use of curvilinear coordinates allows for a deeper insight into a number of important concepts and aspects of, in particular, large strain mechanics problems. These include the notions of the Push Forward operation, Lie derivatives and objective rates.

As will become clear, note that all the tensor relations expressed in symbolic notation already discussed, such as $\mathbf{U} = \sqrt{\mathbf{C}}$, $\mathbf{F}\hat{\mathbf{N}}_i = \lambda_i \mathbf{n}_i$, $\dot{\mathbf{F}} = \mathbf{l}\mathbf{F}$, etc., are independent of coordinate system, and hold also for the convected coordinates discussed here.

2.10.1 Convected Coordinates

In the Cartesian system, orthogonal coordinates X^i, x^i were used. Here, introduce the curvilinear coordinates Θ^i. The material coordinates can then be written as

$$\mathbf{X} = \mathbf{X}(\Theta^1, \Theta^2, \Theta^3) \tag{2.10.1}$$

so $\mathbf{X} = X^i \mathbf{E}_i$ and

$$d\mathbf{X} = dX^i \mathbf{E}_i = d\Theta^i \mathbf{G}_i , \tag{2.10.2}$$

where $\mathbf{G}_i$ are the covariant base vectors in the reference configuration, with corresponding contravariant base vectors $\mathbf{G}^i$, Fig. 2.10.1, with

$$\mathbf{G}^i \cdot \mathbf{G}_j = \delta^i_j \tag{2.10.3}$$

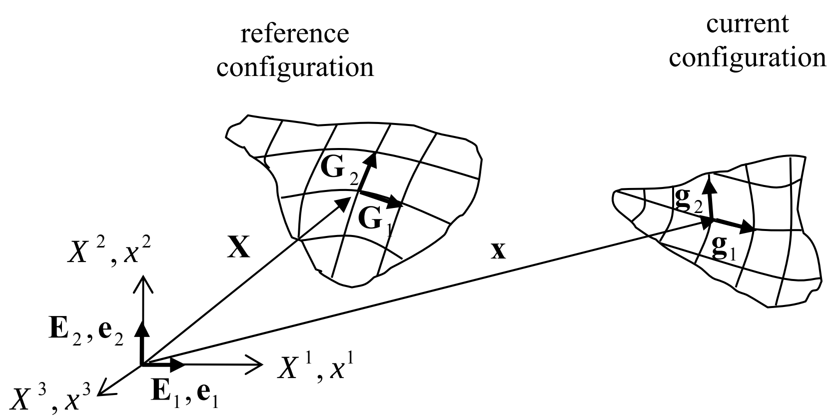


Figure 2.10.1: Curvilinear Coordinates

The coordinate curves form a net in the undeformed configuration (over the surfaces of constant Θ^i). One says that the curvilinear coordinates are **convected** or **embedded**, that is, the coordinate curves are attached to material particles and deform with the body, so that each material particle *has the same values* of the coordinates Θ^i in both the reference and current configurations. The covariant base vectors are tangent the coordinate curves.

In the current configuration, the spatial coordinates can be expressed in terms of a new, "current", set of curvilinear coordinates

$$\mathbf{x} = \mathbf{x}(\Theta^1, \Theta^2, \Theta^3, t), \tag{2.10.4}$$

with corresponding covariant base vectors $\mathbf{g}_i$ and contravariant base vectors $\mathbf{g}^i$, with

$$d\mathbf{x} = dx^i \mathbf{e}_i = d\Theta^i \mathbf{g}_i \quad , \tag{2.10.5}$$

As the material deforms, the covariant base vectors $\mathbf{g}_i$ deform with the body, being "attached" to the body. However, note that the contravariant base vectors $\mathbf{g}^i$ are not as such attached; they have to be re-evaluated at each step of the deformation anew, so as to ensure that the relevant relations, e.g. $\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j$, are always satisfied.

Example 1

Consider a pure shear deformation, where a square deforms into a parallelogram, as illustrated in Fig. 2.10.2. In this scenario, a unit vector $\mathbf{E}_2$ in the "square" gets mapped to a vector $\mathbf{g}_2$ in the parallelogram[1]. The magnitude of $\mathbf{g}_2$ is $1/\sin\alpha$.

[1] This differs from the example worked through in section 1.16; there, the vector $\mathbf{g}_2$ maintained unit magnitude.

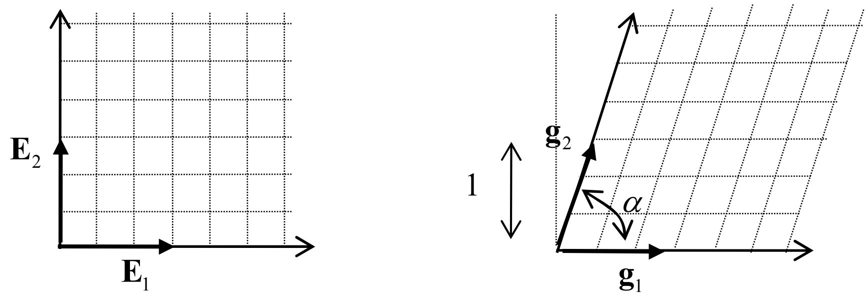


Figure 2.10.2: A pure shear deformation

Consider now a parallelogram (initial condition) deforming into a new parallelogram (the current configuration), as shown in Fig. 2.10.3.

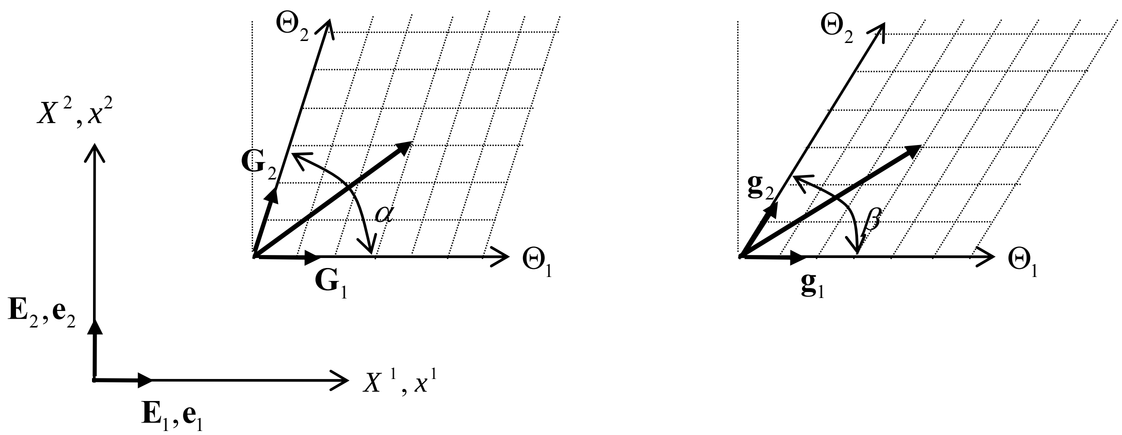


Figure 2.10.3: A pure shear deformation of one parallelogram into another

Keeping in mind that the vector $\mathbf{g}_2$ will be of magnitude $1/\sin\alpha$, the transformation equations 2.10.1 for the configurations shown in Fig. 2.10.3 are[2]

$$\begin{aligned}
&\Theta^1 = X^1 - \frac{1}{\tan\alpha}X^2, \quad \Theta^2 = X^2, \quad \Theta^3 = X^3 \\
&X^1 = \Theta^1 + \frac{1}{\tan\alpha}\Theta^2, \quad X^2 = \Theta^2, \quad X^3 = \Theta^3 \\
&\Theta^1 = x^1 - \frac{1}{\tan\beta}x^2, \quad \Theta^2 = x^2, \quad \Theta^3 = x^3 \\
&x^1 = \Theta^1 + \frac{1}{\tan\beta}\Theta^2, \quad x^2 = \Theta^2, \quad x^3 = \Theta^3
\end{aligned} \tag{2.10.6}$$

[2] Constants have been omitted from these expressions (which represent the translation of the "parallelogram origin" from the Cartesian origin).

Following on from §1.16, Eqns. 1.16.19, the covariant base vectors are:

$$\mathbf{G}_i = \frac{\partial X^m}{\partial \Theta^i}\mathbf{E}_m, \qquad \mathbf{G}_1 = \mathbf{E}_1, \quad \mathbf{G}_2 = \frac{1}{\tan\alpha}\mathbf{E}_1 + \mathbf{E}_2, \quad \mathbf{G}_3 = \mathbf{E}_3$$
$$\mathbf{g}_i = \frac{\partial X^m}{\partial \Theta^i}\mathbf{e}_m, \qquad \mathbf{g}_1 = \mathbf{e}_1, \quad \mathbf{g}_2 = \frac{1}{\tan\beta}\mathbf{e}_1 + \mathbf{e}_2, \quad \mathbf{g}_3 = \mathbf{e}_3 \tag{2.10.7}$$

and the inverse expressions

$$\mathbf{E}_1 = \mathbf{G}_1, \qquad \mathbf{E}_2 = -\frac{1}{\tan\alpha}\mathbf{G}_1 + \mathbf{G}_2, \qquad \mathbf{E}_3 = \mathbf{G}_3$$
$$\mathbf{e}_1 = \mathbf{g}_1, \qquad \mathbf{e}_2 = -\frac{1}{\tan\beta}\mathbf{g}_1 + \mathbf{g}_2, \qquad \mathbf{e}_3 = \mathbf{g}_3 \tag{2.10.8}$$

Line elements in the configurations can now be expressed as

$$d\mathbf{X} = dX^i\mathbf{E}_i = \frac{d\mathbf{X}}{\partial\Theta^i}d\Theta^i = d\Theta^i\mathbf{G}_i$$
$$d\mathbf{x} = dx^i\mathbf{e}_i = \frac{d\mathbf{x}}{\partial\Theta^i}d\Theta^i = d\Theta^i\mathbf{g}_i \tag{2.10.9}$$

The scale factors, i.e. the magnitudes of the covariant base vectors, are (see Eqns. 1.16.36)

$$H_1 = |\mathbf{G}_1| = 1, \quad H_2 = |\mathbf{G}_2| = \frac{1}{\sin\alpha}$$
$$h_1 = |\mathbf{g}_1| = 1, \quad h_2 = |\mathbf{g}_2| = \frac{1}{\sin\beta} \tag{2.10.10}$$

The contravariant base vectors are (see Eqn. 1.16.23)

$$\mathbf{G}^i = \frac{\partial\Theta^i}{\partial X^m}\mathbf{E}_m, \qquad \mathbf{G}^1 = \mathbf{E}_1 - \frac{1}{\tan\alpha}\mathbf{E}_2, \quad \mathbf{G}^2 = \mathbf{E}_2, \quad \mathbf{G}^3 = \mathbf{E}_3$$
$$\mathbf{g}^i = \frac{\partial\Theta^i}{\partial X^m}\mathbf{e}_m, \qquad \mathbf{g}^1 = \mathbf{e}_1 - \frac{1}{\tan\beta}\mathbf{e}_2, \quad \mathbf{g}^2 = \mathbf{e}_2, \quad \mathbf{g}^3 = \mathbf{e}_3 \tag{2.10.11}$$

and the inverse expressions

$$
\begin{aligned}
\mathbf{E}_1 &= \mathbf{G}^1 + \frac{1}{\tan\alpha}\mathbf{G}^2, \qquad \mathbf{E}_2 = \mathbf{G}^2, \qquad \mathbf{E}_3 = \mathbf{G}^3 \\
\mathbf{e}_1 &= \mathbf{g}^1 + \frac{1}{\tan\beta}\mathbf{g}^2, \qquad \mathbf{e}_2 = \mathbf{g}^2, \qquad \mathbf{e}_3 = \mathbf{g}^3
\end{aligned}
\tag{2.10.12}
$$

The magnitudes of the contravariant base vectors, are

$$
\begin{aligned}
H^1 &= \left|\mathbf{G}^1\right| = \frac{1}{\sin\alpha}, \quad H^2 = \left|\mathbf{G}^2\right| = 1 \\
h^1 &= \left|\mathbf{g}^1\right| = \frac{1}{\sin\beta}, \quad h^2 = \left|\mathbf{g}^2\right| = 1
\end{aligned}
\tag{2.10.13}
$$

The metric coefficients are (see Eqns. 1.16.27)

$$
\begin{aligned}
G_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j &= \begin{bmatrix} 1 & \dfrac{1}{\tan\alpha} & 0 \\ \dfrac{1}{\tan\alpha} & \dfrac{1}{\sin^2\alpha} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad G^{ij} = \mathbf{G}^i \cdot \mathbf{G}^j = \begin{bmatrix} \dfrac{1}{\sin^2\alpha} & -\dfrac{1}{\tan\alpha} & 0 \\ -\dfrac{1}{\tan\alpha} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j &= \begin{bmatrix} 1 & \dfrac{1}{\tan\beta} & 0 \\ \dfrac{1}{\tan\beta} & \dfrac{1}{\sin^2\beta} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j = \begin{bmatrix} \dfrac{1}{\sin^2\beta} & -\dfrac{1}{\tan\beta} & 0 \\ -\dfrac{1}{\tan\beta} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
\tag{2.10.14}
$$

The transformation determinants are (consistent with zero volume change), from Eqns. 1.16.32-34,

$$
\begin{aligned}
G &= \det\left[G_{ij}\right] = \frac{1}{\det\left[G^{ij}\right]} = \left(\det\left[\frac{\partial X^i}{\partial \Theta^j}\right]\right)^2 = J_G^2 = 1 \\
g &= \det\left[g_{ij}\right] = \frac{1}{\det\left[g^{ij}\right]} = \left(\det\left[\frac{\partial x^i}{\partial \Theta^j}\right]\right)^2 = J_g^2 = 1
\end{aligned}
\tag{2.10.15}
$$

■

Example 2

Consider a motion whereby a cube of material, with sides of length L_0, is transformed into a cylinder of radius R and height H, Fig. 2.10.4.

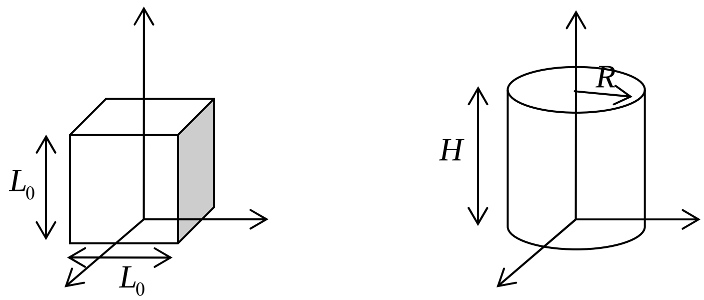


Figure 2.10.4: a cube deformed into a cylinder

A plane view of one quarter of the cube and cylinder are shown in Fig. 2.10.5.

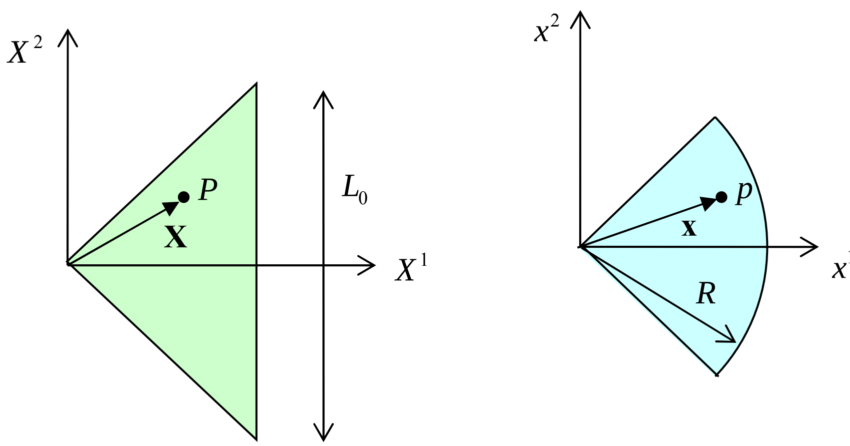


Figure 2.10.5: a cube deformed into a cylinder

The motion and inverse motion are given by

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}), \qquad \begin{aligned} x^1 &= \frac{2R}{L_0} \frac{\left(X^1\right)^2}{\sqrt{\left(X^1\right)^2 + \left(X^2\right)^2}} \\ x^2 &= \frac{2R}{L_0} \frac{X^1 X^2}{\sqrt{\left(X^1\right)^2 + \left(X^2\right)^2}} \\ x^3 &= \frac{H}{L_0} X^3 \end{aligned} \qquad (\text{basis: } \mathbf{e}_i) \tag{2.10.16}$$

and

$$\mathbf{X} = \boldsymbol{\chi}^{-1}(\mathbf{x}), \qquad \begin{aligned} X^1 &= \frac{L_0}{2R}\sqrt{\left(x^1\right)^2 + \left(x^2\right)^2} \\ X^2 &= \frac{L_0}{2R}\frac{x^2}{x^1}\sqrt{\left(x^1\right)^2 + \left(x^2\right)^2} \\ X^3 &= \frac{L_0}{H}x^3 \end{aligned} \qquad (\text{basis: } \mathbf{E}_i) \tag{2.10.17}$$

Introducing a set of convected coordinates, Fig. 2.10.6, the material and spatial coordinates are

$$\mathbf{X} = \mathbf{X}(\Theta^1, \Theta^2, \Theta^3), \qquad \begin{aligned} X^1 &= \left(\frac{L_0}{2R}\right)\Theta^1 \\ X^2 &= \left(\frac{L_0}{2R}\right)\Theta^1 \tan\Theta^2 \\ X^3 &= \frac{L_0}{H}\Theta^3 \end{aligned} \tag{2.10.18}$$

and (these are simply cylindrical coordinates)

$$\mathbf{x} = \mathbf{x}(\Theta^1, \Theta^2, \Theta^3), \qquad \begin{aligned} x^1 &= \Theta^1 \cos\Theta^2 \\ x^2 &= \Theta^1 \sin\Theta^2 \\ x^3 &= \Theta^3 \end{aligned} \tag{2.10.19}$$

A typical material particle (denoted by *p*) is shown in Fig. 2.10.6. Note that the position vectors for *p* have the same Θ^i values, since they represent the same material particle.

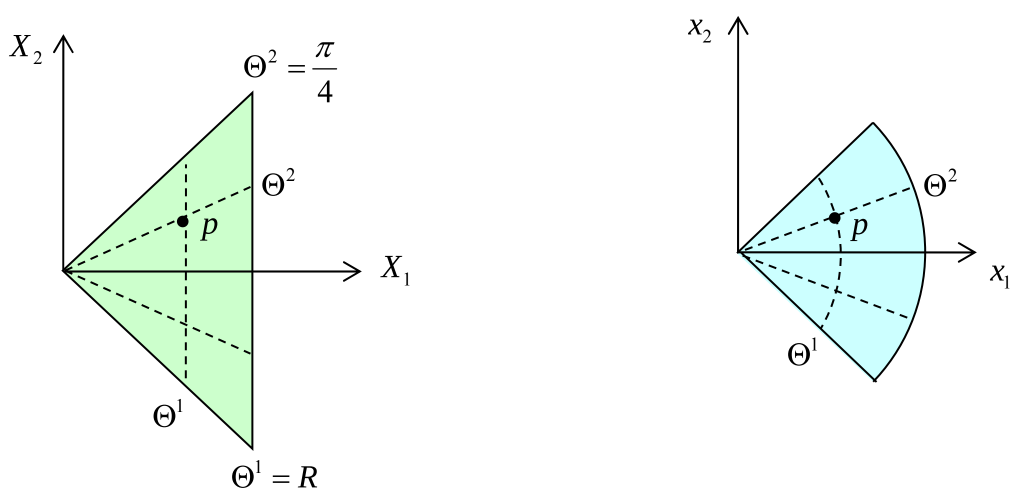


Figure 2.10.6: curvilinear coordinate curves

■

2.10.2 The Deformation Gradient

With convected curvilinear coordinates, the deformation gradient is

$$\begin{aligned}\mathbf{F} &= \mathbf{g}_i \otimes \mathbf{G}^i \\ &= \mathbf{g}_1 \otimes \mathbf{G}^1 + \mathbf{g}_2 \otimes \mathbf{G}^2 + \mathbf{g}_3 \otimes \mathbf{G}^3 , \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\mathbf{g}_i \otimes \mathbf{G}^j\right)\end{aligned} \tag{2.10.20}$$

The deformation gradient operates on a material vector (with contravariant components) $\mathbf{V} = V^i \mathbf{G}_i$, resulting in a spatial tensor $\mathbf{v} = v^i \mathbf{g}_i$ (with the same components $V = v^i$), for example,

$$\mathbf{F} d\mathbf{X} = \left(\mathbf{g}_i \otimes \mathbf{G}^i\right) d\Theta^j \mathbf{G}_j = d\Theta^i \mathbf{g}_i = d\mathbf{x} \tag{2.10.21}$$

To emphasise the point, line elements mapped between the configurations have the same coordinates Θ^i : a line element $d\Theta^1 \mathbf{G}_1 + d\Theta^2 \mathbf{G}_2 + d\Theta^3 \mathbf{G}_3$ gets mapped to

$$\left(\mathbf{g}_1 \otimes \mathbf{G}^1 + \mathbf{g}_2 \otimes \mathbf{G}^2 + \mathbf{g}_3 \otimes \mathbf{G}^3\right)\left(d\Theta^1 \mathbf{G}_1 + d\Theta^2 \mathbf{G}_2 + d\Theta^3 \mathbf{G}_3\right) = d\Theta^1 \mathbf{g}_1 + d\Theta^2 \mathbf{g}_2 + d\Theta^3 \mathbf{g}_3 \tag{2.10.22}$$

This shows also that line elements tangent to the coordinate curves are mapped to new elements tangent to the new coordinate curves; the covariant base vectors $\mathbf{G}_i$ are a field of tangent vectors which get mapped to the new field of tangent vectors $\mathbf{g}_i$, as illustrated in Fig. 2.10.7.

Figure 2.10.7: Vectors tangent to coordinate curves

The deformation gradient **F**, the transpose $\mathbf{F}^{\mathrm{T}}$ and the inverses $\mathbf{F}^{-1}$, $\mathbf{F}^{-\mathrm{T}}$, map the base vectors in one configuration onto the base vectors in the other configuration (that the $\mathbf{F}^{-1}$ and $\mathbf{F}^{-\mathrm{T}}$ in this equation are indeed the inverses of $\mathbf{F}$ and $\mathbf{F}^{\mathrm{T}}$ follows from 1.16.63):

$$\boxed{\begin{aligned} \mathbf{F} &= \mathbf{g}_i \otimes \mathbf{G}^i \\ \mathbf{F}^{-1} &= \mathbf{G}_i \otimes \mathbf{g}^i \\ \mathbf{F}^{-\mathrm{T}} &= \mathbf{g}^i \otimes \mathbf{G}_i \\ \mathbf{F}^{\mathrm{T}} &= \mathbf{G}^i \otimes \mathbf{g}_i \end{aligned}} \quad \rightarrow \quad \boxed{\begin{aligned} \mathbf{F}\mathbf{G}_i &= \mathbf{g}_i \\ \mathbf{F}^{-1}\mathbf{g}_i &= \mathbf{G}_i \\ \mathbf{F}^{-\mathrm{T}}\mathbf{G}^i &= \mathbf{g}^i \\ \mathbf{F}^{\mathrm{T}}\mathbf{g}^i &= \mathbf{G}^i \end{aligned}} \qquad \textbf{Deformation Gradient} \qquad (2.10.23)$$

Thus the tensors $\mathbf{F}$ and $\mathbf{F}^{-1}$ map the covariant base vectors into each other, whereas the tensors $\mathbf{F}^{-\mathrm{T}}$ and $\mathbf{F}^{\mathrm{T}}$ map the contravariant base vectors into each other, as illustrated in Fig. 2.10.8.

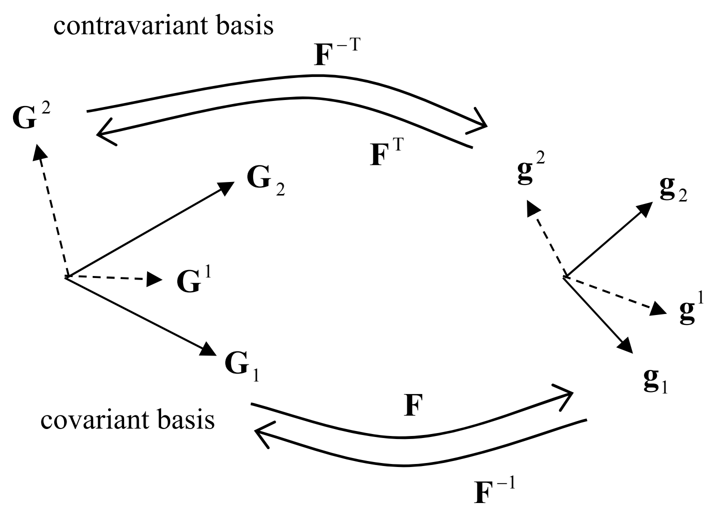


Figure 2.10.8: the deformation gradient, its transpose and the inverses

It was mentioned above how the deformation gradient maps base vectors tangential to the coordinate curves into new vectors tangential to the coordinate curves in the current configuration. In the same way, contravariant base vectors, which are normal to coordinate surfaces, get mapped to normal vectors in the current configuration. For example, the contravariant vector $\mathbf{G}^1$ is normal to the surface of constant Θ^1, and gets mapped through $\mathbf{F}^{-\mathrm{T}}$ to the new vector $\mathbf{g}^1$, which is normal to the surface of constant Θ^1 in the current configuration.

Example 1 continued

Carrying on Example 1 from above, in Cartesian coordinates, 4 corners of an initial parallelogram (see Fig. 2.10.3) get mapped as follows:

$$\begin{gathered}(0,0) \to (0,0) \\ (1,0) \to (1,0) \\ (1/\tan\alpha, 1) \to (1/\tan\beta, 1) \\ (1+1/\tan\alpha, 1) \to (1+1/\tan\beta, 1)\end{gathered} \tag{2.10.24}$$

This corresponds to a deformation gradient with respect to the Cartesian bases:

$$\mathbf{F} = \begin{bmatrix} 1 & \Pi \\ 0 & 1 \end{bmatrix} \quad (\mathbf{E}_i \otimes \mathbf{E}_j), (\mathbf{e}_i \otimes \mathbf{e}_j) \tag{2.10.25}$$

where

$$\Pi = \frac{1}{\tan\beta} - \frac{1}{\tan\alpha} \tag{2.10.26}$$

From the earlier work with example 1, the deformation gradient can be re-expressed in terms of different base vectors:

$$\begin{aligned}\mathbf{F} &= (\mathbf{E}_1 \otimes \mathbf{E}_1) + \Pi(\mathbf{E}_1 \otimes \mathbf{E}_2) + (\mathbf{E}_2 \otimes \mathbf{E}_2) \\ &= (\mathbf{e}_1 \otimes \mathbf{E}_1) + \Pi(\mathbf{e}_1 \otimes \mathbf{E}_2) + (\mathbf{e}_2 \otimes \mathbf{E}_2) \\ &= \mathbf{g}_1 \otimes \left(\mathbf{G}^1 + \frac{1}{\tan\alpha}\mathbf{G}^2\right) + \Pi(\mathbf{g}_1 \otimes \mathbf{G}^2) + \left(-\frac{1}{\tan\beta}\mathbf{g}_1 + \mathbf{g}_2\right) \otimes \mathbf{G}^2 \\ &= \mathbf{g}_i \otimes \mathbf{G}^i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} (\mathbf{g}_i \otimes \mathbf{G}^j)\end{aligned} \tag{2.10.27}$$

which is Eqn. 2.10.20.

In fact, $\mathbf{F}$ can be expressed in a multitude of different ways, depending on which base vectors are used. For example, from the above, $\mathbf{F}$ can also be expressed as

$$\begin{aligned}\mathbf{F} &= (\mathbf{E}_1 \otimes \mathbf{E}_1) + \Pi(\mathbf{E}_1 \otimes \mathbf{E}_2) + (\mathbf{E}_2 \otimes \mathbf{E}_2) \\ &= \left(\mathbf{G}^1 + \frac{1}{\tan\alpha}\mathbf{G}^2\right) \otimes \left(\mathbf{G}^1 + \frac{1}{\tan\alpha}\mathbf{G}^2\right) + \Pi\left[\left(\mathbf{G}^1 + \frac{1}{\tan\alpha}\mathbf{G}^2\right) \otimes \left(\mathbf{G}^2\right)\right] + \left[\mathbf{G}^2 \otimes \mathbf{G}^2\right] \\ &= \begin{bmatrix} 1 & \dfrac{1}{\tan\beta} & 0 \\ \dfrac{1}{\tan\alpha} & \dfrac{1}{\tan\alpha}\dfrac{1}{\tan\beta} & 0 \\ 0 & 0 & 1 \end{bmatrix} (\mathbf{G}^i \otimes \mathbf{G}^j)\end{aligned} \tag{2.10.28}$$

(This can be verified using Eqn. 2.10.30a below.)

Components of F

The various components of **F** and its inverses and the transposes, with respect to the different bases, are:

$$\begin{aligned}
\mathbf{F} &= F_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = F^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = F_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = F_{\cdot j}^{i}\mathbf{G}_i \otimes \mathbf{G}^j \\
&= f_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = f^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = f_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = f_{\cdot j}^{i}\mathbf{g}_i \otimes \mathbf{g}^j \\[1ex]
\mathbf{F}^{-1} &= (F^{-1})_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = (F^{-1})^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = (F^{-1})_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = (F^{-1})_{\cdot j}^{i}\mathbf{G}_i \otimes \mathbf{G}^j \\
&= (f^{-1})_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = (f^{-1})^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = (f^{-1})_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = (f^{-1})_{\cdot j}^{i}\mathbf{g}_i \otimes \mathbf{g}^j \\[1ex]
\mathbf{F}^{\mathrm{T}} &= (F^{\mathrm{T}})_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = (F^{\mathrm{T}})^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = (F^{\mathrm{T}})_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = (F^{\mathrm{T}})_{\cdot j}^{i}\mathbf{G}_i \otimes \mathbf{G}^j \\
&= (f^{\mathrm{T}})_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = (f^{\mathrm{T}})^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = (f^{\mathrm{T}})_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = (f^{\mathrm{T}})_{\cdot j}^{i}\mathbf{g}_i \otimes \mathbf{g}^j \\[1ex]
\mathbf{F}^{-\mathrm{T}} &= (F^{-\mathrm{T}})_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = (F^{-\mathrm{T}})^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = (F^{-\mathrm{T}})_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = (F^{-\mathrm{T}})_{\cdot j}^{i}\mathbf{G}_i \otimes \mathbf{G}^j \\
&= (f^{-\mathrm{T}})_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = (f^{-\mathrm{T}})^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = (f^{-\mathrm{T}})_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = (f^{-\mathrm{T}})_{\cdot j}^{i}\mathbf{g}_i \otimes \mathbf{g}^j
\end{aligned} \tag{2.10.29}$$

The components of **F** with respect to the reference bases $\{\mathbf{G}_i\}, \{\mathbf{G}^i\}$ are

$$
\begin{aligned}
F_{ij} &= \mathbf{G}_i \mathbf{F} \mathbf{G}_j = \mathbf{G}_i \cdot \mathbf{g}_j = \frac{\partial X^m}{\partial \Theta^i} \frac{\partial x^m}{\partial \Theta^j} \\
F^{ij} &= \mathbf{G}^i \mathbf{F} \mathbf{G}^j = G^{jk} \mathbf{G}^i \cdot \mathbf{g}_k \\
F_i^{\cdot j} &= \mathbf{G}_i \mathbf{F} \mathbf{G}^j = G^{jk} \mathbf{G}_i \cdot \mathbf{g}_k \\
F_{\cdot j}^{i} &= \mathbf{G}^i \mathbf{F} \mathbf{G}_j = \mathbf{G}^i \cdot \mathbf{g}_j = \frac{\partial \Theta^i}{\partial X^m} \frac{\partial x^m}{\partial \Theta^j}
\end{aligned}
\tag{2.10.30}
$$

and similarly for the components with respect to the current bases.

Components of the Base Vectors in different Bases

The base vectors themselves can be expressed alternately:

$$
\begin{aligned}
\mathbf{g}_i = \mathbf{F}\mathbf{G}_i &= F_{mj} \left(\mathbf{G}^m \otimes \mathbf{G}^j \right) \mathbf{G}_i &&= F_{\cdot j}^{m} \left(\mathbf{G}_m \otimes \mathbf{G}^j \right) \mathbf{G}_i \\
&= F_{mj} \mathbf{G}^m \delta_i^j &&= F_{\cdot j}^{m} \mathbf{G}_m \delta_i^j \\
&= F_{mi} \mathbf{G}^m &&= F_{\cdot i}^{m} \mathbf{G}_m
\end{aligned}
\tag{2.10.31}
$$

showing that some of the components of the deformation gradient can be viewed also as components of the base vectors. Similarly,

$$
\mathbf{G}_i = \mathbf{F}^{-1} \mathbf{g}_i = \left(f^{-1} \right)_{mi} \mathbf{g}^m = \left(f^{-1} \right)_{\cdot i}^{m} \mathbf{g}_m \tag{2.10.32}
$$

For the contravariant base vectors, one has

$$
\begin{aligned}
\mathbf{g}^i = \mathbf{F}^{-\mathrm{T}} \mathbf{G}^i &= \left(F^{-\mathrm{T}} \right)^{mj} \left(\mathbf{G}_m \otimes \mathbf{G}_j \right) \mathbf{G}^i &&= \left(F^{-\mathrm{T}} \right)_{m}^{\cdot j} \left(\mathbf{G}^m \otimes \mathbf{G}_j \right) \mathbf{G}^i \\
&= \left(F^{-\mathrm{T}} \right)^{mj} \mathbf{G}_m \delta_j^i &&= \left(F^{-\mathrm{T}} \right)_{m}^{\cdot j} \mathbf{G}^m \delta_j^i \\
&= \left(F^{-\mathrm{T}} \right)^{mi} \mathbf{G}_m &&= \left(F^{-\mathrm{T}} \right)_{m}^{\cdot i} \mathbf{G}^m
\end{aligned}
\tag{2.10.33}
$$

and

$$
\mathbf{G}^i = \mathbf{F}^{\mathrm{T}} \mathbf{g}^i = \left(f^{\mathrm{T}} \right)^{mi} \mathbf{g}_m = \left(f^{\mathrm{T}} \right)_{m}^{\cdot i} \mathbf{g}^m \tag{2.10.34}
$$

2.10.3 Reduction to Material and Spatial Coordinates

Material Coordinates

Suppose that the material coordinates X^i with Cartesian basis are used (rather than the convected coordinates with curvilinear basis $\mathbf{G}_i$), Fig. 2.10.9. Then

$$\Theta^i \to X^i, \qquad \begin{aligned} \mathbf{G}_i &= \frac{\partial X^j}{\partial \Theta^i}\mathbf{E}_j = \frac{\partial X^j}{\partial X^i}\mathbf{E}_j = \mathbf{E}_i \\ \mathbf{G}^i &= \frac{\partial \Theta^i}{\partial X^j}\mathbf{E}^j = \frac{\partial X^i}{\partial X^j}\mathbf{E}^j = \mathbf{E}^i \end{aligned}, \qquad \begin{aligned} \mathbf{g}_i &= \frac{\partial x^j}{\partial \Theta^i}\mathbf{e}_j = \frac{\partial x^j}{\partial X^i}\mathbf{e}_j \\ \mathbf{g}^i &= \frac{\partial \Theta^i}{\partial x^j}\mathbf{e}^j = \frac{\partial X^i}{\partial x^j}\mathbf{e}^j \end{aligned} \tag{2.10.35}$$

and

$$\begin{aligned} \mathbf{F} &= \mathbf{g}_i \otimes \mathbf{G}^i = \mathbf{g}_i \otimes \mathbf{E}^i = \frac{\partial x^j}{\partial X^i}\mathbf{e}_j \otimes \mathbf{E}^i = \mathrm{Grad}\mathbf{x} \\ \mathbf{F}^{-1} &= \mathbf{G}_i \otimes \mathbf{g}^i = \mathbf{E}_i \otimes \mathbf{g}^i = \frac{\partial X^i}{\partial x^j}\mathbf{E}_i \otimes \mathbf{e}^j = \mathrm{grad}\mathbf{X} \end{aligned} \tag{2.10.36}$$

which are Eqns. 2.2.2, 2.2.4. Thus $\mathrm{Grad}\mathbf{x}$ is the notation for $\mathbf{F}$ and $\mathrm{grad}\mathbf{X}$ is the notation for $\mathbf{F}^{-1}$, to be used when the material coordinates X_i are used to describe the deformation.

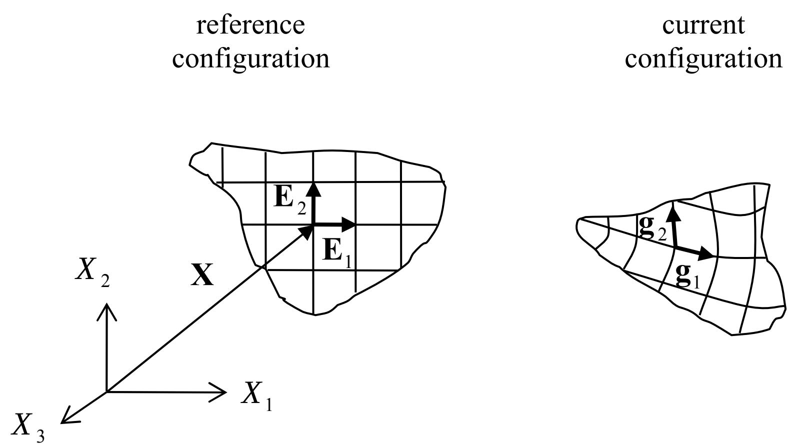


Figure 2.10.9: Material coordinates and deformed basis

Spatial Coordinates

Similarly, when the spatial coordinates x^i are to be used as independent variables, then

$$\Theta^i \to x^i, \qquad \begin{aligned} \mathbf{G}_i &= \frac{\partial X^j}{\partial \Theta^i}\mathbf{E}_j = \frac{\partial X^j}{\partial x^i}\mathbf{E}_j \\ \mathbf{G}^i &= \frac{\partial \Theta^i}{\partial X^j}\mathbf{E}^j = \frac{\partial x^i}{\partial X^j}\mathbf{E}^j \end{aligned}, \qquad \begin{aligned} \mathbf{g}_i &= \frac{\partial x^j}{\partial \Theta^i}\mathbf{e}_j = \frac{\partial x^j}{\partial x^i}\mathbf{e}_j = \mathbf{e}_i \\ \mathbf{g}^i &= \frac{\partial \Theta^i}{\partial x^j}\mathbf{e}^j = \frac{\partial x^i}{\partial x^j}\mathbf{e}^j = \mathbf{e}^i \end{aligned} \tag{2.10.37}$$

and

$$\mathbf{F} = \mathbf{g}_i \otimes \mathbf{G}^i = \mathbf{e}_i \otimes \mathbf{G}^i = \frac{\partial x^i}{\partial X^j}\mathbf{e}_i \otimes \mathbf{E}^j = \mathrm{Grad}\mathbf{x}$$
$$\mathbf{F}^{-1} = \mathbf{G}_i \otimes \mathbf{g}^i = \mathbf{G}_i \otimes \mathbf{e}^i = \frac{\partial X^j}{\partial x^i}\mathbf{E}_j \otimes \mathbf{e}^i = \mathrm{grad}\mathbf{X} \tag{2.10.38}$$

The descriptions are illustrated in Fig. 2.10.10. Note that the base vectors $\mathbf{G}_i$, $\mathbf{g}_i$ are not the same in each of these cases (curvilinear, material and spatial).

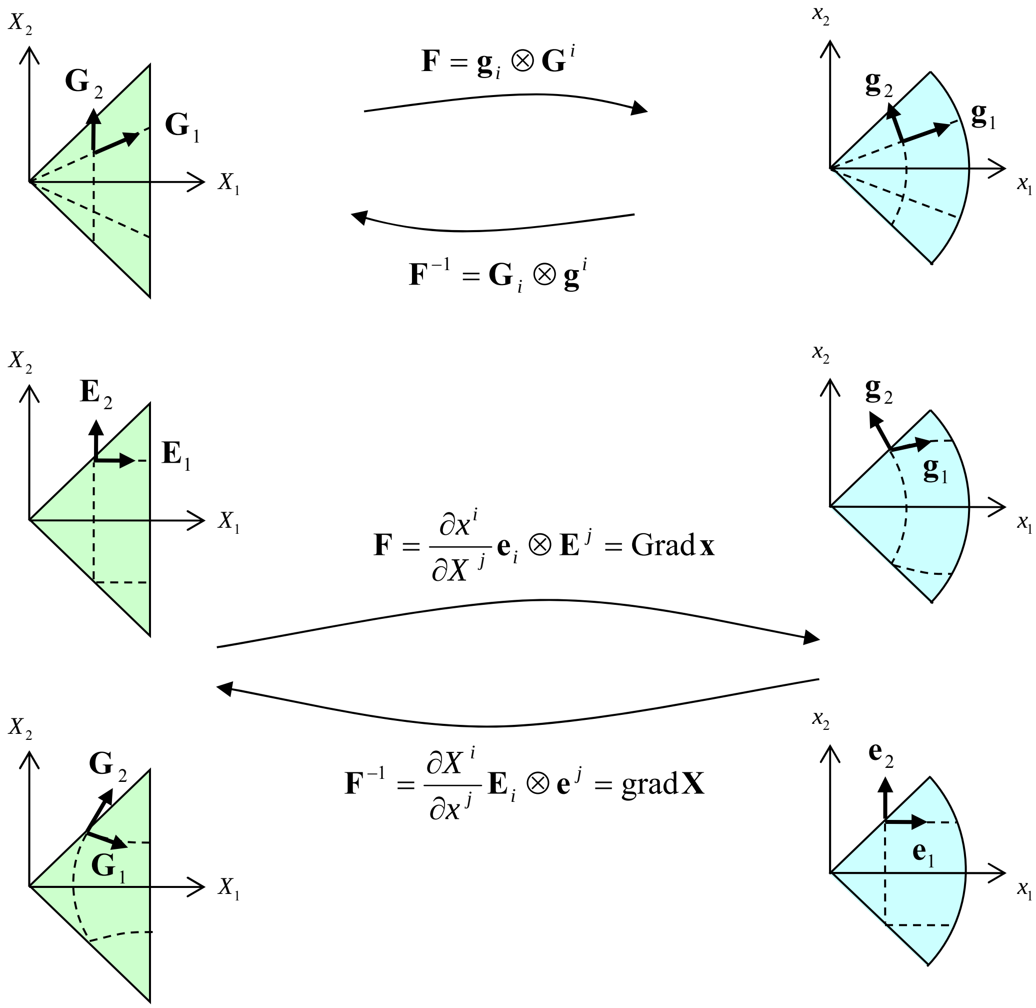


Figure 2.10.10: deformation described using different independent variables

2.10.4 Strain Tensors

The Cauchy-Green tensors

The right Cauchy-Green tensor **C** and the left Cauchy-Green tensor **b** are defined by Eqns. 2.2.10, 2.2.13,

$$\begin{aligned}
\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} &= \left(\mathbf{G}^i \otimes \mathbf{g}_i\right)\left(\mathbf{g}_j \otimes \mathbf{G}^j\right) = g_{ij}\mathbf{G}^i \otimes \mathbf{G}^j \equiv C_{ij}\mathbf{G}^i \otimes \mathbf{G}^j \\
\mathbf{C}^{-1} = \mathbf{F}^{-1}\mathbf{F}^{-\mathrm{T}} &= \left(\mathbf{G}_i \otimes \mathbf{g}^i\right)\left(\mathbf{g}^j \otimes \mathbf{G}_j\right) = g^{ij}\mathbf{G}_i \otimes \mathbf{G}_j \equiv \left(C^{-1}\right)^{ij}\mathbf{G}_i \otimes \mathbf{G}_j \\
\mathbf{b} = \mathbf{F}\mathbf{F}^{\mathrm{T}} &= \left(\mathbf{g}_i \otimes \mathbf{G}^i\right)\left(\mathbf{G}^j \otimes \mathbf{g}_j\right) = G^{ij}\mathbf{g}_i \otimes \mathbf{g}_j \equiv b^{ij}\mathbf{g}_i \otimes \mathbf{g}_j \\
\mathbf{b}^{-1} = \mathbf{F}^{-\mathrm{T}}\mathbf{F}^{-1} &= \left(\mathbf{g}^i \otimes \mathbf{G}_i\right)\left(\mathbf{G}_j \otimes \mathbf{g}^j\right) = G_{ij}\mathbf{g}^i \otimes \mathbf{g}^j \equiv \left(b^{-1}\right)_{ij}\mathbf{g}^i \otimes \mathbf{g}^j
\end{aligned} \tag{2.10.39}$$

Thus the covariant components of the right Cauchy-Green tensor are the metric coefficients g_{ij}. This highlights the importance of **C**: the $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$ give a clear measure of the deformation occurring. (It is possible to evaluate other components of **C**, e.g. C^{ij}, and also its components with respect to the current basis, but only the components C_{ij} with respect to the reference basis are (normally) used in the analysis.)

The Stretch

Now, analogous to 2.2.9, 2.2.12,

$$\begin{aligned}
ds^2 &= d\mathbf{x} \cdot d\mathbf{x} = d\mathbf{X}\mathbf{C}d\mathbf{X} \\
dS^2 &= d\mathbf{X} \cdot d\mathbf{X} = d\mathbf{x}\mathbf{b}^{-1}d\mathbf{x}
\end{aligned} \tag{2.10.40}$$

so that the stretches are, analogous to 2.2.17,

$$\begin{aligned}
\lambda^2 &= \frac{ds^2}{dS^2} = \frac{d\mathbf{X}}{|d\mathbf{X}|}\mathbf{C}\frac{d\mathbf{X}}{|d\mathbf{X}|} = d\hat{\mathbf{X}}\mathbf{C}d\hat{\mathbf{X}} \quad \to d\hat{X}^i C_{ij} d\hat{X}^j \\
\frac{1}{\lambda^2} &= \frac{dS^2}{ds^2} = \frac{d\mathbf{x}}{|d\mathbf{x}|}\mathbf{b}^{-1}\frac{d\mathbf{x}}{|d\mathbf{x}|} = d\hat{\mathbf{x}}\mathbf{b}^{-1}d\hat{\mathbf{x}} \quad \to d\hat{x}^i \left(b^{-1}\right)_{ij} d\hat{x}^j
\end{aligned} \tag{2.10.41}$$

The Green-Lagrange and Euler-Almansi Tensors

The Green-Lagrange strain tensor **E** and the Euler-Almansi strain tensor **e** are defined through 2.2.22, 2.2.24,

$$\frac{ds^2 - dS^2}{2} = d\mathbf{X}\frac{1}{2}(\mathbf{C} - \mathbf{I})d\mathbf{X} \equiv d\mathbf{X}\mathbf{E}d\mathbf{X}$$
$$\frac{ds^2 - dS^2}{2} = d\mathbf{x}\frac{1}{2}(\mathbf{I} - \mathbf{b}^{-1})d\mathbf{x} \equiv d\mathbf{x}\mathbf{e}d\mathbf{x} \tag{2.10.42}$$

The components of **E** and **e** can be evaluated through (writing $\mathbf{G} \equiv \mathbf{I}$, the identity tensor expressed in terms of the base vectors in the reference configuration, and $\mathbf{g} \equiv \mathbf{I}$, the identity tensor expressed in terms of the base vectors in the current configuration)

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{G}) = \frac{1}{2}\left(g_{ij}\mathbf{G}^i \otimes \mathbf{G}^j - G_{ij}\mathbf{G}^i \otimes \mathbf{G}^j\right) = \frac{1}{2}\left(g_{ij} - G_{ij}\right)\mathbf{G}^i \otimes \mathbf{G}^j \equiv E_{ij}\mathbf{G}^i \otimes \mathbf{G}^j$$
$$\mathbf{e} = \frac{1}{2}\left(\mathbf{g} - \mathbf{b}^{-1}\right) = \frac{1}{2}\left(g_{ij}\mathbf{g}^i \otimes \mathbf{g}^j - G_{ij}\mathbf{g}^i \otimes \mathbf{g}^j\right) = \frac{1}{2}\left(g_{ij} - G_{ij}\right)\mathbf{g}^i \otimes \mathbf{g}^j \equiv e_{ij}\mathbf{g}^i \otimes \mathbf{g}^j \tag{2.10.43}$$

Note that the components of **E** and **e** with respect to their bases are equal, $E_{ij} = e_{ij}$ (although this is not true regarding their other components, e.g. $E^{ij} \neq e^{ij}$).

Example 1 continued

Carrying on Example 1 from above, consider now an example vector

$$\mathbf{V} = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad (\mathbf{E}_i) \tag{2.10.44}$$

The contravariant and covariant components are

$$\mathbf{V} = \begin{bmatrix} V_x - \dfrac{1}{\tan\alpha}V_y \\ V_y \end{bmatrix} \quad (\mathbf{G}_i), \qquad \mathbf{V} = \begin{bmatrix} V_x \\ \dfrac{1}{\tan\alpha}V_x + V_y \end{bmatrix} \quad (\mathbf{G}^i) \tag{2.10.45}$$

The magnitude of the vector can be calculated through (see Eqn. 1.16.52 and 1.16.49)

$$\begin{aligned}
|\mathbf{V}| &= \sqrt{\overset{\mathbf{E}_i}{\mathbf{V}} \cdot \overset{\mathbf{E}_i}{\mathbf{V}}} = \sqrt{V_x^2 + V_y^2} \\
&= \sqrt{\overset{\mathbf{G}_i}{\mathbf{V}} \cdot \overset{\mathbf{G}_i}{\mathbf{V}}} = \sqrt{G_{ij}V^iV^j} = \sqrt{\left(V_x - \frac{V_y}{\tan\alpha}\right)^2 G_{11} + 2\left(V_x - \frac{V_y}{\tan\alpha}\right)V_yG_{12} + V_y^2G_{22}} \\
&= \sqrt{\overset{\mathbf{G}^i}{\mathbf{V}} \cdot \overset{\mathbf{G}^i}{\mathbf{V}}} = \sqrt{G^{ij}V_iV_j} = \sqrt{V_x^2G^{11} + 2V_x\left(\frac{V_x}{\tan\alpha} + V_y\right)G^{12} + \left(\frac{V_x}{\tan\alpha} + V_y\right)^2 G^{22}}
\end{aligned} \tag{2.10.46}$$

The new vector is obtained from the deformation gradient:

$$\begin{aligned}\mathbf{v} = \mathbf{F}\overset{\mathbf{E}_i}{\mathbf{V}} &= \begin{bmatrix} 1 & \Pi \\ 0 & 1 \end{bmatrix}\begin{bmatrix} V_x \\ V_y \end{bmatrix} &&= \begin{bmatrix} V_x + \Pi V_y \\ V_y \end{bmatrix}(\mathbf{e}_i) \\ = \mathbf{F}\overset{\mathbf{G}_i}{\mathbf{V}} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} V_x - \dfrac{V_y}{\tan\alpha} \\ V_y \end{bmatrix} &&= \begin{bmatrix} V_x - \dfrac{1}{\tan\alpha}V_y \\ V_y \end{bmatrix}(\mathbf{g}_i)\end{aligned} \tag{2.10.47}$$

In terms of the contravariant vectors:

$$\mathbf{v} = v_j g^j = \begin{bmatrix} V_x + \Pi V_y \\ \dfrac{1}{\tan\beta}V_x + \left(1 + \dfrac{1}{\tan\beta}\Pi\right)V_y \end{bmatrix}(\mathbf{g}^i) \tag{2.10.48}$$

Note that the contravariant components do not change with the deformation, but the covariant components do in general change with the deformation.

The magnitudes of the vectors before and after deformation are given by the Cauchy-Green strain tensors, whose coefficients are those of the metric tensors (the first of these is the same as 2.10.46)

$$\begin{aligned}\mathbf{V}\cdot\mathbf{V} &= \mathbf{F}^{-1}\overset{\mathbf{g}_i}{\mathbf{v}}\cdot\mathbf{F}^{-1}\overset{\mathbf{g}_i}{\mathbf{v}} = \overset{\mathbf{g}_i}{\mathbf{v}}\mathbf{F}^{-\mathrm{T}}\mathbf{F}^{-1}\overset{\mathbf{g}_i}{\mathbf{v}} = \overset{\mathbf{g}_i}{\mathbf{v}}\mathbf{b}^{-1}\overset{\mathbf{g}_i}{\mathbf{v}} = v^k\mathbf{g}_k G_{ij}\mathbf{g}^i \otimes \mathbf{g}^j v^l \mathbf{g}_l = G_{ij}v^i v^j \\ \mathbf{v}\cdot\mathbf{v} &= \mathbf{F}\overset{\mathbf{G}_i}{\mathbf{V}}\cdot\mathbf{F}\overset{\mathbf{G}_i}{\mathbf{V}} = \overset{\mathbf{G}_i}{\mathbf{V}}\mathbf{F}^{\mathrm{T}}\mathbf{F}\overset{\mathbf{G}_i}{\mathbf{V}} = \overset{\mathbf{G}_i}{\mathbf{V}}\mathbf{C}\overset{\mathbf{G}_i}{\mathbf{V}} = V^k\mathbf{G}_k g_{ij}\mathbf{G}^i \otimes \mathbf{G}^j V^l \mathbf{G}_l = g_{ij}V^i V^j\end{aligned} \tag{2.10.49}$$

From this, the magnitude of the vector after deformation is

$$\sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{g_{ij}V^iV^j} = \sqrt{\left(V_x^2 + V_y^2\right) + \Pi V_y\left(2V_x + \Pi V_y\right)} \tag{2.10.50}$$

2.10.5 Intermediate Configurations

Stretch and Rotation Tensors

The polar decompositions $\mathbf{F} = \mathbf{RU} = \mathbf{vR}$ have been described in §2.2.5. The decompositions are illustrated in Fig. 2.10.11. In the material decomposition, the material is first stretched by $\mathbf{U}$ and then rotated by $\mathbf{R}$. Let the base vectors in the associated intermediate configuration be $\{\hat{\mathbf{g}}_i\}$. Similarly, in the spatial decomposition, the material is first rotated by $\mathbf{R}$ and then stretched by $\mathbf{v}$. Let the base vectors in the associated intermediate configuration in this case be $\{\mathbf{G}_i\}$. Then, analogous to Eqn. 2.10.23, {▲Problem 1}

$$\begin{aligned} \mathbf{U} &= \hat{\mathbf{g}}_i \otimes \mathbf{G}^i \\ \mathbf{U}^{-1} &= \mathbf{G}_i \otimes \hat{\mathbf{g}}^i \\ \mathbf{U}^{-\mathrm{T}} &= \hat{\mathbf{g}}^i \otimes \mathbf{G}_i \\ \mathbf{U}^{\mathrm{T}} &= \mathbf{G}^i \otimes \hat{\mathbf{g}}_i \end{aligned} \quad \rightarrow \quad \begin{aligned} \mathbf{U}\mathbf{G}_i &= \hat{\mathbf{g}}_i \\ \mathbf{U}^{-1}\hat{\mathbf{g}}_i &= \mathbf{G}_i \\ \mathbf{U}^{-\mathrm{T}}\mathbf{G}^i &= \hat{\mathbf{g}}^i \\ \mathbf{U}^{\mathrm{T}}\hat{\mathbf{g}}^i &= \mathbf{G}^i \end{aligned} \tag{2.10.51}$$

$$\begin{aligned} \mathbf{v} &= \mathbf{g}_i \otimes \hat{\mathbf{G}}^i \\ \mathbf{v}^{-1} &= \hat{\mathbf{G}}_i \otimes \mathbf{g}^i \\ \mathbf{v}^{-\mathrm{T}} &= \mathbf{g}^i \otimes \hat{\mathbf{G}}_i \\ \mathbf{v}^{\mathrm{T}} &= \hat{\mathbf{G}}^i \otimes \mathbf{g}_i \end{aligned} \quad \rightarrow \quad \begin{aligned} \mathbf{v}\hat{\mathbf{G}}_i &= \mathbf{g}_i \\ \mathbf{v}^{-1}\mathbf{g}_i &= \hat{\mathbf{G}}_i \\ \mathbf{v}^{-\mathrm{T}}\hat{\mathbf{G}}^i &= \mathbf{g}^i \\ \mathbf{v}^{\mathrm{T}}\mathbf{g}^i &= \hat{\mathbf{G}}^i \end{aligned} \tag{2.10.52}$$

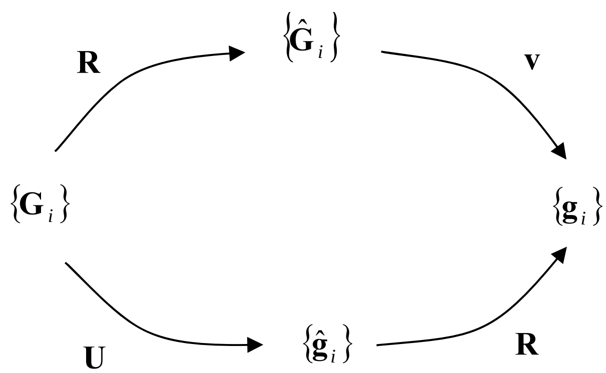


Figure 2.10.11: the material and spatial polar decompositions

Note that **U** and **v** symmetric, $\mathbf{U} = \mathbf{U}^{\mathrm{T}}$, $\mathbf{v} = \mathbf{v}^{\mathrm{T}}$, so

$$\begin{aligned} \mathbf{U} &= \hat{\mathbf{g}}_i \otimes \mathbf{G}^i = \mathbf{G}^i \otimes \hat{\mathbf{g}}_i \\ \mathbf{U}^{-1} &= \mathbf{G}_i \otimes \hat{\mathbf{g}}^i = \hat{\mathbf{g}}^i \otimes \mathbf{G}_i \end{aligned} \quad \rightarrow \quad \begin{aligned} \mathbf{U}\mathbf{G}_i &= \hat{\mathbf{g}}_i, \quad \mathbf{U}\hat{\mathbf{g}}^i = \mathbf{G}^i \\ \mathbf{U}^{-1}\hat{\mathbf{g}}_i &= \mathbf{G}_i, \quad \mathbf{U}^{-1}\mathbf{G}^i = \hat{\mathbf{g}}^i \end{aligned} \tag{2.10.53}$$

$$\begin{aligned} \mathbf{v} &= \mathbf{g}_i \otimes \hat{\mathbf{G}}^i = \hat{\mathbf{G}}^i \otimes \mathbf{g}_i \\ \mathbf{v}^{-1} &= \hat{\mathbf{G}}_i \otimes \mathbf{g}^i = \mathbf{g}^i \otimes \hat{\mathbf{G}}_i \end{aligned} \quad \rightarrow \quad \begin{aligned} \mathbf{v}\hat{\mathbf{G}}_i &= \mathbf{g}_i, \quad \mathbf{v}\mathbf{g}^i = \hat{\mathbf{G}}^i \\ \mathbf{v}^{-1}\mathbf{g}_i &= \hat{\mathbf{G}}_i, \quad \mathbf{v}^{-1}\hat{\mathbf{G}}^i = \mathbf{g}^i \end{aligned} \tag{2.10.54}$$

Similarly, for the rotation tensor, with **R** orthogonal, $\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}}$,

$$\begin{aligned} \mathbf{R} &= \hat{\mathbf{G}}_i \otimes \mathbf{G}^i = \hat{\mathbf{G}}^i \otimes \mathbf{G}_i \\ \mathbf{R}^{\mathrm{T}} &= \mathbf{G}_i \otimes \hat{\mathbf{G}}^i = \mathbf{G}^i \otimes \hat{\mathbf{G}}_i \end{aligned} \quad \rightarrow \quad \begin{aligned} \mathbf{R}\mathbf{G}_i &= \hat{\mathbf{G}}_i, \quad \mathbf{R}\mathbf{G}^i = \hat{\mathbf{G}}^i \\ \mathbf{R}^{\mathrm{T}}\hat{\mathbf{G}}_i &= \mathbf{G}_i, \quad \mathbf{R}^{\mathrm{T}}\hat{\mathbf{G}}^i = \mathbf{G}^i \end{aligned} \tag{2.10.55}$$

$$\begin{aligned} \mathbf{R} &= \mathbf{g}_i \otimes \hat{\mathbf{g}}^i = \mathbf{g}^i \otimes \hat{\mathbf{g}}_i \\ \mathbf{R}^{\mathrm{T}} &= \hat{\mathbf{g}}_i \otimes \mathbf{g}^i = \hat{\mathbf{g}}^i \otimes \mathbf{g}_i \end{aligned} \quad \rightarrow \quad \begin{aligned} \mathbf{R}\hat{\mathbf{g}}_i &= \mathbf{g}_i, \quad \mathbf{R}\hat{\mathbf{g}}^i = \mathbf{g}^i \\ \mathbf{R}^{\mathrm{T}}\mathbf{g}_i &= \hat{\mathbf{g}}_i, \quad \mathbf{R}^{\mathrm{T}}\mathbf{g}^i = \hat{\mathbf{g}}^i \end{aligned} \tag{2.10.56}$$

The above relations can be checked using Eqns. 2.10.23 and $\mathbf{F} = \mathbf{RU}$, $\mathbf{F} = \mathbf{vR}$, $\mathbf{v}^{-1} = \mathbf{RF}^{-1}$, etc.

Various relations between the base vectors can be derived, for example,

$$\begin{aligned}
\hat{\mathbf{G}}_i \cdot \mathbf{g}_j &= (\mathbf{RG}_i)\cdot(\mathbf{R}\hat{\mathbf{g}}_j) = \mathbf{G}_i\mathbf{R}^{\mathrm{T}}\mathbf{R}\hat{\mathbf{g}}_j &&= \mathbf{G}_i \cdot \hat{\mathbf{g}}_j \\
\hat{\mathbf{G}}^i \cdot \mathbf{g}^j &= \qquad\qquad \cdots &&= \mathbf{G}^i \cdot \hat{\mathbf{g}}^j \\
\hat{\mathbf{G}}^i \cdot \mathbf{g}_j &= \qquad\qquad \cdots &&= \mathbf{G}^i \cdot \hat{\mathbf{g}}_j \\
\hat{\mathbf{G}}_i \cdot \mathbf{g}^j &= \qquad\qquad \cdots &&= \mathbf{G}_i \cdot \hat{\mathbf{g}}^j
\end{aligned} \tag{2.10.57}$$

Deformation Gradient Relationship between Bases

The various base vectors are related above through the stretch and rotation tensors. The intermediate bases are related directly through the deformation gradient. For example, from 2.10.53a, 2.10.55b,

$$\hat{\mathbf{g}}_i = \mathbf{UG}_i = \mathbf{UR}^{\mathrm{T}}\hat{\mathbf{G}}_i = \mathbf{F}^{\mathrm{T}}\hat{\mathbf{G}}_i \tag{2.10.58}$$

In the same way,

$$\begin{aligned}
\hat{\mathbf{g}}_i &= \mathbf{F}^{\mathrm{T}}\hat{\mathbf{G}}_i \\
\hat{\mathbf{g}}^i &= \mathbf{F}^{-1}\hat{\mathbf{G}}^i \\
\hat{\mathbf{G}}_i &= \mathbf{F}^{-\mathrm{T}}\hat{\mathbf{g}}_i \\
\hat{\mathbf{G}}^i &= \mathbf{F}\hat{\mathbf{g}}^i
\end{aligned} \tag{2.10.59}$$

Tensor Components

The stretch and rotation tensors can be decomposed along any of the bases. For $\mathbf{U}$ the most natural bases would be $\{\mathbf{G}_i\}$ and $\{\mathbf{G}^i\}$, for example,

$$\begin{aligned}
\mathbf{U} &= U_{ij}\mathbf{G}^i \otimes \mathbf{G}^j, \quad & U_{ij} &= \mathbf{G}_i\mathbf{UG}_j = \mathbf{G}_i \cdot \hat{\mathbf{g}}_j \\
\mathbf{U} &= U^{ij}\mathbf{G}_i \otimes \mathbf{G}_j, \quad & U^{ij} &= \mathbf{G}^i\mathbf{UG}^j = G^{im}\mathbf{G}^j \cdot \hat{\mathbf{g}}_m \\
\mathbf{U} &= U^{i}_{\cdot j}\mathbf{G}_i \otimes \mathbf{G}^j, \quad & U^{i}_{\cdot j} &= \mathbf{G}^i\mathbf{UG}_j = \mathbf{G}^i \cdot \hat{\mathbf{g}}_j \\
\mathbf{U} &= U_{i}^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j, \quad & U_{i}^{\cdot j} &= \mathbf{G}_i\mathbf{UG}^j = \hat{\mathbf{g}}_i \cdot \mathbf{G}^j
\end{aligned} \tag{2.10.60}$$

with $U_{ij} = U_{ji}, U^{ij} = U^{ji}, U^{i}_{\cdot j} = U_{j}^{\cdot i}, U_{i}^{\cdot j} = U^{j}_{\cdot i}$. One also has

$$\begin{aligned}
\mathbf{v} &= v_{ij}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}^j, \quad v_{ij} = \hat{\mathbf{G}}_i \mathbf{v} \hat{\mathbf{G}}_j = \hat{\mathbf{G}}_i \cdot \mathbf{g}_j \\
\mathbf{v} &= v^{ij}\hat{\mathbf{G}}_i \otimes \hat{\mathbf{G}}_j, \quad v^{ij} = \hat{\mathbf{G}}^i \mathbf{v} \hat{\mathbf{G}}^j = \hat{G}^{im}\hat{\mathbf{G}}^j \cdot \mathbf{g}_m \\
\mathbf{v} &= v^i_{\cdot j}\hat{\mathbf{G}}_i \otimes \hat{\mathbf{G}}^j, \quad v^i_{\cdot j} = \hat{\mathbf{G}}^i \mathbf{v} \hat{\mathbf{G}}_j = \hat{\mathbf{G}}^i \cdot \mathbf{g}_j \\
\mathbf{v} &= v_i^{\cdot j}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}_j, \quad v_i^{\cdot j} = \hat{\mathbf{G}}_i \mathbf{v} \hat{\mathbf{G}}^j = \mathbf{g}_i \cdot \hat{\mathbf{G}}^j
\end{aligned} \tag{2.10.61}$$

with similar symmetry. Also,

$$\begin{aligned}
\mathbf{U}^{-1} &= \left(U^{-1}\right)_{ij}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}^j, \quad \left(U^{-1}\right)_{ij} = \hat{\mathbf{g}}_i \mathbf{U}^{-1} \hat{\mathbf{g}}_j = \mathbf{G}_i \cdot \hat{\mathbf{g}}_j \\
\mathbf{U}^{-1} &= \left(U^{-1}\right)^{ij}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}_j, \quad \left(U^{-1}\right)^{ij} = \hat{\mathbf{g}}^i \mathbf{U}^{-1} \hat{\mathbf{g}}^j = \hat{g}^{im}\mathbf{G}_m \cdot \hat{\mathbf{g}}^j \\
\mathbf{U}^{-1} &= \left(U^{-1}\right)^i_{\cdot j}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}^j, \quad \left(U^{-1}\right)^i_{\cdot j} = \hat{\mathbf{g}}^i \mathbf{U}^{-1} \hat{\mathbf{g}}_j = \hat{\mathbf{g}}^i \cdot \mathbf{G}_j \\
\mathbf{U}^{-1} &= \left(U^{-1}\right)_i^{\cdot j}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}_j, \quad \left(U^{-1}\right)_i^{\cdot j} = \hat{\mathbf{g}}_i \mathbf{U}^{-1} \hat{\mathbf{g}}^j = \mathbf{G}_i \cdot \hat{\mathbf{g}}^j
\end{aligned} \tag{2.10.62}$$

and

$$\begin{aligned}
\mathbf{v}^{-1} &= \left(v^{-1}\right)_{ij}\mathbf{g}^i \otimes \mathbf{g}^j, \quad \left(v^{-1}\right)_{ij} = \mathbf{g}_i \mathbf{v}^{-1} \mathbf{g}_j = \hat{\mathbf{G}}_i \cdot \mathbf{g}_j \\
\mathbf{v}^{-1} &= \left(v^{-1}\right)^{ij}\mathbf{g}_i \otimes \mathbf{g}_j, \quad \left(v^{-1}\right)^{ij} = \mathbf{g}^i \mathbf{v}^{-1} \mathbf{g}^j = g^{mj}\hat{\mathbf{G}}_m \cdot \mathbf{g}^i \\
\mathbf{v}^{-1} &= \left(v^{-1}\right)^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j, \quad \left(v^{-1}\right)^i_{\cdot j} = \mathbf{g}^i \mathbf{v}^{-1} \mathbf{g}_j = \mathbf{g}^i \cdot \hat{\mathbf{G}}_j \\
\mathbf{v}^{-1} &= \left(v^{-1}\right)_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j, \quad \left(v^{-1}\right)_i^{\cdot j} = \mathbf{g}_i \mathbf{v}^{-1} \mathbf{g}^j = \hat{\mathbf{G}}_i \cdot \mathbf{g}^j
\end{aligned} \tag{2.10.63}$$

with similar symmetry. Note that, comparing 2.10.60a, 2.10.61a, 2.10.62a, 2.10.63a and using 2.10.57,

$$\begin{aligned}
\mathbf{U} &= U_{ij}\mathbf{G}^i \otimes \mathbf{G}^j \\
\mathbf{v} &= v_{ij}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}^j \\
\mathbf{U}^{-1} &= \left(U^{-1}\right)_{ij}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}^j \\
\mathbf{v}^{-1} &= \left(v^{-1}\right)_{ij}\mathbf{g}^i \otimes \mathbf{g}^j
\end{aligned} \qquad U_{ij} = \left(U^{-1}\right)_{ij} = v_{ij} = \left(v^{-1}\right)_{ij} \tag{2.10.64}$$

Now note that rotations preserve vectors lengths and, in particular, preserve the metric, i.e.,

$$\begin{aligned}
G_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j \quad &= \quad \hat{G}_{ij} = \hat{\mathbf{G}}_i \cdot \hat{\mathbf{G}}_j \\
g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j \quad &= \quad \hat{g}_{ij} = \hat{\mathbf{g}}_i \cdot \hat{\mathbf{g}}_j
\end{aligned} \tag{2.10.65}$$

Thus, again using 2.10.57, and 2.10.60-2.10.63, the contravariant components of the above tensors are also equal, $U^{ij} = \left(U^{-1}\right)^{ij} = v^{ij} = \left(v^{-1}\right)^{ij}$.

As mentioned, the tensors can be decomposed along other bases, for example,

$$\mathbf{v} = v^{ij}\mathbf{g}_i \otimes \mathbf{g}_j, \quad v^{ij} = \mathbf{g}^i \mathbf{v} \mathbf{g}^j = \hat{\mathbf{G}}^i \cdot \mathbf{g}^j \tag{2.10.66}$$

2.10.6 Eigenvectors and Eigenvalues

Analogous to §2.2.5, the eigenvalues of **C** are determined from the eigenvalue problem

$$\det\left(\mathbf{C} - \lambda_{\mathbf{C}}\mathbf{I}\right) = 0 \tag{2.10.67}$$

leading to the characteristic equation 1.11.5

$$\lambda_{\mathbf{C}}^3 - \mathrm{I}_{\mathbf{C}}\lambda_{\mathbf{C}}^2 + \mathrm{II}_{\mathbf{C}}\lambda_{\mathbf{C}} - \mathrm{III}_{\mathbf{C}} = 0 \tag{2.10.68}$$

with principal scalar invariants 1.11.6-7

$$\begin{aligned} \mathrm{I}_{\mathbf{C}} &= \mathrm{tr}\mathbf{C} = A_i^i = \lambda_{\mathbf{C}1} + \lambda_{\mathbf{C}2} + \lambda_{\mathbf{C}3} \\ \mathrm{II}_{\mathbf{C}} &= \tfrac{1}{2}\left[(\mathrm{tr}\mathbf{C})^2 - \mathrm{tr}(\mathbf{C}^2)\right] = \tfrac{1}{2}\left(C_i^i C_j^j - C_j^i C_i^j\right) = \lambda_{\mathbf{C}1}\lambda_{\mathbf{C}2} + \lambda_{\mathbf{C}2}\lambda_{\mathbf{C}3} + \lambda_{\mathbf{C}3}\lambda_{\mathbf{C}1} \\ \mathrm{III}_{\mathbf{C}} &= \det\mathbf{C} = \varepsilon_{ijk} C_1^i C_2^j C_3^k = \lambda_{\mathbf{C}1}\lambda_{\mathbf{C}2}\lambda_{\mathbf{C}3} \end{aligned} \tag{2.10.69}$$

The eigenvectors are the principal material directions $\hat{\mathbf{N}}_i$, with

$$\left(\mathbf{C} - \lambda_i \mathbf{I}\right)\hat{\mathbf{N}}_i = \mathbf{0} \tag{2.10.70}$$

The spectral decomposition is then

$$\mathbf{C} = \sum_{i=1}^{3} \lambda_i^2 \hat{\mathbf{N}}_i \otimes \hat{\mathbf{N}}_i \tag{2.10.71}$$

where $\lambda_{\mathbf{C}i} = \lambda_i^2$ and the λ_i are the stretches. The remaining spectral decompositions in 2.2.37 hold also. Note also that the rotation tensor in terms of principal directions is (see 2.2.35)

$$\mathbf{R} = \hat{\mathbf{n}}_i \otimes \hat{\mathbf{N}}^i = \hat{\mathbf{n}}^i \otimes \hat{\mathbf{N}}_i \tag{2.10.72}$$

where $\hat{\mathbf{n}}_i$ are the spatial principal directions.

2.10.7 Displacement and Displacement Gradients

Consider the displacement $\mathbf{u}$ of a material particle. This can be written in terms of covariant components U_i and u_i:

$$\mathbf{u} = \mathbf{x} - \mathbf{X} \equiv U_i \mathbf{G}^i = u_i \mathbf{g}^i . \tag{2.10.73}$$

The covariant derivative of $\mathbf{u}$ can be expressed as

$$\frac{\partial \mathbf{u}}{\partial \Theta^i} = U_m\big|_i \mathbf{G}^m = u_m\big\|_i \mathbf{g}^m \tag{2.10.74}$$

The single line refers to covariant differentiation with respect to the undeformed basis, i.e. the Christoffel symbols to use are functions of the G_{ij}. The double line refers to covariant differentiation with respect to the deformed basis, i.e. the Christoffel symbols to use are functions of the g_{ij}.

Alternatively, the covariant derivative can be expressed as

$$\frac{\partial \mathbf{u}}{\partial \Theta^i} = \frac{\partial \mathbf{x}}{\partial \Theta^i} - \frac{\partial \mathbf{X}}{\partial \Theta^i} = \mathbf{g}_i - \mathbf{G}_i \tag{2.10.75}$$

and so

$$\begin{aligned} \mathbf{g}_i &= \mathbf{G}_i + U_m\big|_i \mathbf{G}^m = \left(\delta_i^m + U^m\big|_i\right)\mathbf{G}_m = F^{m}_{\cdot i}\mathbf{G}_m \\ \mathbf{G}_i &= \mathbf{g}_i - u_m\big\|_i \mathbf{g}^m = \left(\delta_i^m - u^m\big\|_i\right)\mathbf{g}_m = \left(f^{-1}\right)^m_{\cdot i}\mathbf{g}_m \end{aligned} \tag{2.10.76}$$

The last equalities following from 2.10.31-32.

The components of the Green-Lagrange and Euler-Almansi strain tensors 2.10.43 can be written in terms of displacements using relations 2.10.76 {▲Problem 2}:

$$\begin{aligned} E_{ij} &= \frac{1}{2}\left(g_{ij} - G_{ij}\right) = \frac{1}{2}\left(U_i\big|_j + U_j\big|_i + U_n\big|_i U^n\big|_j\right) \\ e_{ij} &= \frac{1}{2}\left(g_{ij} - G_{ij}\right) = \frac{1}{2}\left(u_i\big\|_j + u_j\big\|_i - u_n\big\|_i u^n\big\|_j\right) \end{aligned} \tag{2.10.77}$$

In terms of spatial coordinates, $\Theta^i = X^i$, $\mathbf{G}_i = \mathbf{E}_i$, $\mathbf{g}_i = \left(\partial x^j / \partial X^i\right)\mathbf{e}_j$, $U_i\big|_j = \partial U_i / \partial X^j$, the components of the Euler-Lagrange strain tensor are

$$E_{ij} = \frac{1}{2}\left(g_{ij} - G_{ij}\right) = \frac{1}{2}\left(\frac{\partial x^m}{\partial X^i}\frac{\partial x^n}{\partial X^j}\delta_{mn} - \delta_{ij}\right) = \frac{1}{2}\left(\frac{\partial U_i}{\partial X^j} + \frac{\partial U_j}{\partial X^i} + \frac{\partial U_k}{\partial X^i}\frac{\partial U_k}{\partial X^j}\right) \quad (2.10.78)$$

which is 2.2.46.

2.10.8 The Deformation of Area and Volume Elements

Differential Volume Element

Consider a differential volume element formed by the elements $d\Theta^i \mathbf{G}_i$ in the undeformed configuration, Eqn. 1.16.43:

$$dV = \sqrt{G}\, d\Theta^1 d\Theta^2 d\Theta^3 \quad (2.10.79)$$

where, Eqn. 1.16.31,1.16.34,

$$\sqrt{G} = \sqrt{\det\left[G_{ij}\right]}, \quad G_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j \quad (2.10.80)$$

The *same* volume element in the deformed configuration is determined by the elements $d\Theta^i \mathbf{g}_i$:

$$dv = \sqrt{g}\, d\Theta^1 d\Theta^2 d\Theta^3 \quad (2.10.81)$$

where

$$\sqrt{g} = \sqrt{\det\left[g_{ij}\right]}, \quad g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j \quad (2.8.82)$$

From 1.16.53 *et seq.*, 2.10.11,

$$\begin{aligned}\sqrt{g} &= \mathbf{g}_1 \cdot \mathbf{g}_2 \times \mathbf{g}_3 \\ &= F^i_{\cdot 1} F^j_{\cdot 2} F^k_{\cdot 3} \mathbf{G}_i \cdot \mathbf{G}_j \times \mathbf{G}_k \\ &= F^i_{\cdot 1} F^j_{\cdot 2} F^k_{\cdot 3} \varepsilon_{ijk} \sqrt{G} \\ &= \sqrt{G} \det \mathbf{F}\end{aligned} \quad (2.10.83)$$

where ε_{ijk} is the Cartesian permutation symbol, and so the Jacobian determinant is (see 2.2.53)

$$J = \frac{dv}{dV} = \frac{\sqrt{g}}{\sqrt{G}} = \det \mathbf{F} \quad (2.10.84)$$

and $\det \mathbf{F}$ is the determinant of the matrix with components $F_{.j}^{i}$.

Differential Area Element

Consider a differential surface (parallelogram) element in the undeformed configuration, bounded by two vector elements $d\mathbf{X}^{(1)}$ and $d\mathbf{X}^{(2)}$, and with unit normal $\hat{\mathbf{N}}$. Then the vector normal to the surface element and with magnitude equal to the area of the surface is, using 1.16.54, given by

$$\hat{\mathbf{N}}dS = d\mathbf{X}^{(1)} \times d\mathbf{X}^{(2)} = d\Theta^{(1)i}\mathbf{G}_i \times d\Theta^{(2)j}\mathbf{G}_j = e_{ijk}^{(\mathbf{G})} d\Theta^{(1)i} d\Theta^{(2)j}\mathbf{G}^k \qquad (2.10.85)$$

where $e_{ijk}^{(\mathbf{G})}$ is the permutation symbol associated with the basis $\mathbf{G}_i$, i.e.

$$e_{ijk}^{(\mathbf{G})} = \varepsilon_{ijk}\mathbf{G}_i \cdot \mathbf{G}_j \times \mathbf{G}_k = \varepsilon_{ijk}\sqrt{G}\,. \qquad (2.10.86)$$

Using $\mathbf{G}^k = \mathbf{F}^{\mathrm{T}}\mathbf{g}^k$, one has

$$\hat{\mathbf{N}}dS = \varepsilon_{ijk}\sqrt{G}\, d\Theta^{(1)i} d\Theta^{(2)j}\mathbf{F}^{\mathrm{T}}\mathbf{g}^k \qquad (2.10.87)$$

Similarly, the surface vector in the deformed configuration with unit normal $\hat{\mathbf{n}}$ is

$$\hat{\mathbf{n}}ds = d\mathbf{x}^{(1)} \times d\mathbf{x}^{(2)} = d\Theta^{(1)i}\mathbf{g}_i \times d\Theta^{(2)j}\mathbf{g}_j = e_{ijk}^{(\mathbf{g})} d\Theta^{(1)i} d\Theta^{(2)j}\mathbf{g}^k \qquad (2.10.88)$$

where $e_{ijk}^{(\mathbf{g})}$ is the permutation symbol associated with the basis $\mathbf{g}_i$, i.e.

$$e_{ijk}^{(\mathbf{g})} = \varepsilon_{ijk}\mathbf{g}_i \cdot \mathbf{g}_j \times \mathbf{g}_k = \varepsilon_{ijk}\sqrt{g}\,. \qquad (2.10.89)$$

Comparing the two expressions for the areas in the undeformed and deformed configurations, 2.10.87-88, one finds that

$$\hat{\mathbf{n}}ds = \sqrt{\frac{g}{G}}\mathbf{F}^{-\mathrm{T}}\mathbf{N}dS = (\det \mathbf{F})\mathbf{F}^{-\mathrm{T}}\mathbf{N}dS \qquad (2.10.90)$$

which is Nanson's relation, Eqn. 2.2.59. This is consistent with was said earlier in relation to Fig. 2.10.8 and the contravariant bases: $\mathbf{F}^{-\mathrm{T}}$ maps vectors normal to the coordinate curves in the initial configuration into corresponding vectors normal to the coordinate curves in the current configuration.

2.10.9 Problems

1. Derive the relations 2.10.51.
2. Use relations 2.10.76, with $g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j$ and $G_{ij} = \mathbf{G}_i \cdot \mathbf{G}_j$, to derive 2.10.77

$$E_{ij} = \frac{1}{2}\left(g_{ij} - G_{ij}\right) = \frac{1}{2}\left(U_i\big|_j + U_j\big|_i + U_n\big|_i U^n\big|_j\right)$$

$$e_{ij} = \frac{1}{2}\left(g_{ij} - G_{ij}\right) = \frac{1}{2}\left(u_i\big\|_j + u_j\big\|_i - u_n\big\|_i u^n\big\|_j\right)$$

Convected Coordinates: Time Rates of Change

In this section, the time derivatives of kinematic tensors described in §2.4-2.6 are now described using convected coordinates.

2.11.1 Deformation Rates

Time Derivatives of the Base Vectors and the Deformation Gradient

The material time derivatives of the material base vectors are zero: $\dot{\mathbf{G}}_i = \dot{\mathbf{G}}^i = 0$. The material time derivatives of the deformed base vectors are, from 2.10.23, (and using $\dot{\mathbf{I}} = d\left(\mathbf{F}\mathbf{F}^{-1}\right)/dt = \dot{\mathbf{F}}\mathbf{F}^{-1} + \mathbf{F}\dot{\mathbf{F}}^{-1}$)

$$\begin{aligned} \dot{\mathbf{g}}_i &= \dot{\mathbf{F}}\mathbf{G}_i = \dot{\mathbf{F}}\mathbf{F}^{-1}\mathbf{g}_i = -\mathbf{F}\dot{\mathbf{F}}^{-1}\mathbf{g}_i \\ \dot{\mathbf{g}}^i &= \dot{\mathbf{F}}^{-\mathrm{T}}\mathbf{G}^i = \dot{\mathbf{F}}^{-\mathrm{T}}\mathbf{F}^{\mathrm{T}}\mathbf{g}^i = -\mathbf{F}^{-\mathrm{T}}\dot{\mathbf{F}}^{\mathrm{T}}\mathbf{g}^i \end{aligned} \tag{2.11.1}$$

with, again from 2.10.23,

$$\begin{aligned} \dot{\mathbf{F}} &= \dot{\mathbf{g}}_i \otimes \mathbf{G}^i \\ \dot{\mathbf{F}}^{-1} &= \mathbf{G}_i \otimes \dot{\mathbf{g}}^i \\ \dot{\mathbf{F}}^{-\mathrm{T}} &= \dot{\mathbf{g}}^i \otimes \mathbf{G}_i \\ \dot{\mathbf{F}}^{\mathrm{T}} &= \mathbf{G}^i \otimes \dot{\mathbf{g}}_i \end{aligned} \tag{2.11.2}$$

The Velocity Gradient

The velocity gradient is defined by 2.5.2, $\mathbf{l} = \operatorname{grad}\mathbf{v}$, so that, using 1.16.23,

$$\mathbf{l} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \frac{\partial \mathbf{v}}{\partial x^i} \otimes \mathbf{e}^i = \frac{\partial \mathbf{v}}{\partial \Theta^j}\frac{\partial \Theta^j}{\partial x^i} \otimes \mathbf{e}^i = \frac{\partial \mathbf{v}}{\partial \Theta^j} \otimes \mathbf{g}^j \tag{2.11.3}$$

Also, from 1.16.19,

$$\dot{\mathbf{g}}_i = \frac{\partial \dot{\mathbf{x}}}{\partial \Theta^i} = \frac{\partial \mathbf{v}}{\partial \Theta^i} \tag{2.11.4}$$

so that, as an alternative to 2.11.3,

$$\mathbf{l} = \dot{\mathbf{g}}_i \otimes \mathbf{g}^i \tag{2.11.5}$$

The components of the spatial velocity gradient are

$$
\begin{aligned}
l_{ij} &= \mathbf{g}_i \mathbf{l} \mathbf{g}_j = \mathbf{g}_i \cdot \dot{\mathbf{g}}_j \\
l^{i}_{\cdot j} &= \mathbf{g}^i \mathbf{l} \mathbf{g}_j = \mathbf{g}^i \cdot \dot{\mathbf{g}}_j \\
l_i^{\cdot j} &= \mathbf{g}_i \mathbf{l} \mathbf{g}^j = g^{mj} \mathbf{g}_i \cdot \dot{\mathbf{g}}_m = \mathbf{g}_i \cdot \dot{\mathbf{g}}^j \\
l^{ij} &= \mathbf{g}^i \mathbf{l} \mathbf{g}^j = \mathbf{g}^i \cdot \dot{\mathbf{g}}^j
\end{aligned}
\tag{2.11.6}
$$

Convected Bases

From 2.11.1, 2.11.2 and 2.11.5,

$$
\begin{aligned}
\dot{\mathbf{g}}_i &= \mathbf{l}\mathbf{g}_i & \dot{\mathbf{g}}^i &= -\mathbf{l}^{\mathrm{T}} \mathbf{g}^i \\
&= \mathbf{g}_i \mathbf{l}^{\mathrm{T}} & &= -\mathbf{g}^i \mathbf{l}
\end{aligned}
\tag{2.11.7}
$$

Contracting the first of these with $d\Theta^i$ leads to

$$
\dot{\mathbf{g}}_i d\Theta^i = \mathbf{l}\mathbf{g}_i d\Theta^i \tag{2.11.8}
$$

which is equivalent to 2.5.1, $d\mathbf{v} = \mathbf{l} d\mathbf{x}$.

Time Derivatives of the Deformation Gradient in terms of the Velocity Gradient

Eqns. 2.11.2 can also be re-expressed using Eqns. 2.11.7:

$$
\begin{aligned}
\dot{\mathbf{F}} &= \dot{\mathbf{g}}_i \otimes \mathbf{G}^i = \mathbf{g}_i \mathbf{l}^{\mathrm{T}} \otimes \mathbf{G}^i = \mathbf{l}\mathbf{g}_i \otimes \mathbf{G}^i = \mathbf{l}\mathbf{F} \\
\dot{\mathbf{F}}^{-1} &= \mathbf{G}_i \otimes \dot{\mathbf{g}}^i = -\mathbf{G}_i \otimes \mathbf{g}^i \mathbf{l} = -\mathbf{F}^{-1} \mathbf{l} \\
\dot{\mathbf{F}}^{-\mathrm{T}} &= \dot{\mathbf{g}}^i \otimes \mathbf{G}_i = -\mathbf{g}^i \mathbf{l} \otimes \mathbf{G}_i = -\mathbf{l}^{\mathrm{T}} \mathbf{g}^i \otimes \mathbf{G}_i = -\mathbf{l}^{\mathrm{T}} \mathbf{F}^{-\mathrm{T}} \\
\dot{\mathbf{F}}^{\mathrm{T}} &= \mathbf{G}^i \otimes \dot{\mathbf{g}}_i = \mathbf{G}^i \otimes \mathbf{g}_i \mathbf{l}^{\mathrm{T}} = \mathbf{F}^{\mathrm{T}} \mathbf{l}^{\mathrm{T}}
\end{aligned}
\tag{2.11.9}
$$

which are Eqns. 2.5.4-5.

An alternative way of arriving at Eqns. 2.11.7 is to start with Eqns. 2.11.9: the covariant base vectors $\mathbf{G}_i$ convect to $\mathbf{g}_i(t)$ over time through the time-dependent deformation gradient: $\mathbf{g}_i(t) = \mathbf{F}(t)\mathbf{G}_i$. For this relation to hold at all times, one must have, from Eqn. 2.11.9b,

$$
\begin{aligned}
\dot{\mathbf{G}}_i = 0 &= \dot{\overline{\mathbf{F}^{-1}\mathbf{g}_i}} \\
&= \dot{\mathbf{F}}^{-1}\mathbf{g}_i + \mathbf{F}^{-1}\dot{\mathbf{g}}_i \\
&= \mathbf{F}^{-1}\left(-\mathbf{l}\mathbf{g}_i + \dot{\mathbf{g}}_i\right)
\end{aligned}
\tag{2.11.10}
$$

Thus, in order to maintain the convection of the tangent basis over time, one requires that

$$\dot{\mathbf{g}}_i = \mathbf{l}\mathbf{g}_i \tag{2.11.11}$$

The contravariant base vectors $\mathbf{G}^i$ transform to $\mathbf{g}^i(t)$ over time through the time-dependent inverse transpose of the deformation gradient: $\mathbf{g}^i(t) = \mathbf{F}^{-\mathrm{T}}(t)\mathbf{G}^i$. For this relation to hold at all times, one must have, from Eqn. 2.11.9d,

$$\begin{aligned}\dot{\mathbf{G}}^i = 0 &= \overline{\mathbf{F}^{\mathrm{T}}\mathbf{g}^i}^{\bullet} \\ &= \dot{\mathbf{F}}^{\mathrm{T}}\mathbf{g}^i + \mathbf{F}^{\mathrm{T}}\dot{\mathbf{g}}^i \\ &= \mathbf{F}^{\mathrm{T}}\left(\mathbf{l}^{\mathrm{T}}\mathbf{g}^i + \dot{\mathbf{g}}^i\right)\end{aligned} \tag{2.11.12}$$

Thus, in order to maintain the convection of the normal basis over time, one requires that

$$\dot{\mathbf{g}}^i = -\mathbf{l}^{\mathrm{T}}\mathbf{g}^i \tag{2.11.13}$$

The Rate of Deformation and Spin Tensors

From 2.5.6, $\mathbf{l} = \mathbf{d} + \mathbf{w}$. The covariant components of the rate of deformation and spin are

$$\begin{aligned} d_{ij} &= \frac{1}{2}\mathbf{g}_i\left(\mathbf{l} + \mathbf{l}^{\mathrm{T}}\right)\mathbf{g}_j = \frac{1}{2}\mathbf{g}_i\left(\dot{\mathbf{g}}_m \otimes \mathbf{g}^m + \mathbf{g}^m \otimes \dot{\mathbf{g}}_m\right)\mathbf{g}_j = \frac{1}{2}\left(\mathbf{g}_i \cdot \dot{\mathbf{g}}_j + \dot{\mathbf{g}}_i \cdot \mathbf{g}_j\right) = \frac{1}{2}\overline{\mathbf{g}_i \cdot \mathbf{g}_j}^{\bullet} \\ w_{ij} &= \frac{1}{2}\mathbf{g}_i\left(\mathbf{l} - \mathbf{l}^{\mathrm{T}}\right)\mathbf{g}_j = \frac{1}{2}\mathbf{g}_i\left(\dot{\mathbf{g}}_m \otimes \mathbf{g}^m - \mathbf{g}^m \otimes \dot{\mathbf{g}}_m\right)\mathbf{g}_j = \frac{1}{2}\left(\mathbf{g}_i \cdot \dot{\mathbf{g}}_j - \dot{\mathbf{g}}_i \cdot \mathbf{g}_j\right)\end{aligned} \tag{2.11.14}$$

Alternatively, from 2.11.6a,

$$\begin{aligned}\mathbf{d} = \frac{1}{2}\left(\mathbf{l} + \mathbf{l}^{\mathrm{T}}\right) &= \frac{1}{2}\left(\mathbf{g}_i \cdot \dot{\mathbf{g}}_j + \dot{\mathbf{g}}_i \cdot \mathbf{g}_j\right)\mathbf{g}_i \otimes \mathbf{g}_j \\ &= \frac{1}{2}\overline{\mathbf{g}_i \cdot \mathbf{g}_j}^{\bullet}\,\mathbf{g}_i \otimes \mathbf{g}_j \\ &= \frac{1}{2}\dot{g}_{ij}\mathbf{g}_i \otimes \mathbf{g}_j\end{aligned} \tag{2.11.15}$$

2.12 Pull Back, Push Forward and Lie Time Derivatives

This section is in the main concerned with the following issue: an observer attached to a fixed, say Cartesian, coordinate system will see a material move and deform over time, and will observe various vectorial and tensorial quantities to change also. However, a hypothetical observer attached to the deforming material, and moving and deforming with the material, will see something different. The question is: what quantities will be seen to change from this embedded observer's viewpoint?

2.12.1 Time Derivatives of Spatial Fields

In terms of the spatial basis, a spatial vector **v** can be expressed in terms of the covariant components and contravariant components,

$$\mathbf{v} = v_i \mathbf{g}^i, \qquad \mathbf{v} = v^i \mathbf{g}_i \tag{2.12.1}$$

We want to distinguish between two quantities. The first is the material time derivative of the vector **v**:

$$\dot{\mathbf{v}} = \dot{\overline{v_i \mathbf{g}^i}} = \dot{v}_i \mathbf{g}^i + v_i \dot{\mathbf{g}}^i, \qquad \dot{\mathbf{v}} = \dot{\overline{v^i \mathbf{g}_i}} = \dot{v}^i \mathbf{g}_i + v^i \dot{\mathbf{g}}_i \tag{2.12.2}$$

The second is the time derivative *holding the base vectors fixed*,

$$\dot{v}_i \mathbf{g}^i, \qquad \dot{v}^i \mathbf{g}_i \tag{2.12.3}$$

This latter is called the **convected derivative** and is the rate of the change as seen by an observer attached to the deforming bases.

From Eqn. 2.12.1, the components of **v** can be expressed as

$$v_i = \mathbf{v} \cdot \mathbf{g}_i, \quad v^i = \mathbf{v} \cdot \mathbf{g}^i \tag{2.12.4}$$

Taking the material time derivative, and using Eqns. 2.11.11, 2.11.13,

$$\begin{aligned} \dot{v}_i &= \dot{\overline{\mathbf{v} \cdot \mathbf{g}_i}} & \qquad \dot{v}^i &= \dot{\overline{\mathbf{v} \cdot \mathbf{g}^i}} \\ &= \dot{\mathbf{v}} \cdot \mathbf{g}_i + \mathbf{v} \cdot \dot{\mathbf{g}}_i \,, & &= \dot{\mathbf{v}} \cdot \mathbf{g}^i + \mathbf{v} \cdot \dot{\mathbf{g}}^i \\ &= \left(\dot{\mathbf{v}} + \mathbf{l}^{\mathrm{T}} \mathbf{v} \right) \cdot \mathbf{g}_i & &= \left(\dot{\mathbf{v}} - \mathbf{l}\mathbf{v} \right) \cdot \mathbf{g}^i \end{aligned} \tag{2.12.5}$$

Thus there are two convected derivatives of a vector, depending on whether one is using covariant or contravariant components:

$$\begin{aligned}\dot{v}_i \mathbf{g}^i &= \dot{\mathbf{v}} + \mathbf{l}^{\mathrm{T}} \mathbf{v} \\ \dot{v}^i \mathbf{g}_i &= \dot{\mathbf{v}} - \mathbf{l}\mathbf{v}\end{aligned} \tag{2.12.6}$$

As will be seen below, these quantities are **Lie derivatives** of the vector $\mathbf{v}$.

The time derivative of the components can be expressed in an alternative way, by expressing the spatial base vectors $\mathbf{g}_i, \mathbf{g}^i$ in terms of the material base vectors $\mathbf{G}_i, \mathbf{G}^i$; using Eqns. 2.10.23:

$$\begin{aligned}\dot{v}_i &= \dot{\overline{\mathbf{v} \cdot \mathbf{g}_i}} & \dot{v}^i &= \dot{\overline{\mathbf{v} \cdot \mathbf{g}^i}} \\ &= \dot{\overline{\mathbf{v} \cdot \mathbf{F}\mathbf{G}_i}}, & &= \dot{\overline{\mathbf{v} \cdot \mathbf{F}^{-\mathrm{T}}\mathbf{G}^i}} \\ &= \dot{\overline{\mathbf{F}^{\mathrm{T}}\mathbf{v}}}\mathbf{G}_i & &= \dot{\overline{\mathbf{F}^{-1}\mathbf{v}}}\mathbf{G}^i\end{aligned} \tag{2.12.7}$$

So, as an alternative to Eqns. 2.12.6,

$$\begin{aligned}\dot{v}_i \mathbf{G}^i &= \dot{\overline{\mathbf{F}^{\mathrm{T}}\mathbf{v}}} \\ \dot{v}^i \mathbf{G}_i &= \dot{\overline{\mathbf{F}^{-1}\mathbf{v}}}\end{aligned} \tag{2.12.8}$$

As will be seen further below, the quantities on the right are the material time derivatives of the **pull-back** of the vector $\mathbf{v}$.

Repeating the above, now for a spatial tensor $\mathbf{a}$: in terms of the spatial basis, $\mathbf{a}$ can be expressed in terms of the covariant components and contravariant components as

$$\mathbf{a} = a_{ij}\mathbf{g}^i \otimes \mathbf{g}^j, \qquad \mathbf{a} = a^{ij}\mathbf{g}_i \otimes \mathbf{g}_j \tag{2.12.9}$$

The material time derivative of the tensor $\mathbf{a}$ is

$$\begin{aligned}\dot{\mathbf{a}} &= \dot{\overline{a_{ij}\mathbf{g}^i \otimes \mathbf{g}^j}} = \dot{a}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j + a_{ij}\dot{\mathbf{g}}^i \otimes \mathbf{g}^j + a_{ij}\mathbf{g}^i \otimes \dot{\mathbf{g}}^j \\ &= \dot{\overline{a^{ij}\mathbf{g}_i \otimes \mathbf{g}_j}} = \dot{a}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j + a^{ij}\dot{\mathbf{g}}_i \otimes \mathbf{g}_j + a^{ij}\mathbf{g}_i \otimes \dot{\mathbf{g}}_j\end{aligned} \tag{2.12.10}$$

and the convected derivative is the first term:

$$\dot{a}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j, \qquad \dot{a}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j \tag{2.12.11}$$

The components of $\mathbf{a}$ can be expressed as

$$a_{ij} = \mathbf{g}_i \mathbf{A} \mathbf{g}_j, \quad a^{ij} = \mathbf{g}^i \mathbf{A} \mathbf{g}^j \tag{2.12.12}$$

Taking the material time derivative, and again using Eqns. 2.11.11, 2.11.13,

$$\begin{aligned} \dot{a}_{ij} &= \dot{\overline{\mathbf{g}_i \mathbf{a} \mathbf{g}_j}} \\ &= \dot{\mathbf{g}}_i \mathbf{a} \mathbf{g}_j + \mathbf{g}_i \dot{\mathbf{a}} \mathbf{g}_j + \mathbf{g}_i \mathbf{a} \dot{\mathbf{g}}_j \,, \\ &= \mathbf{g}_i \left(\dot{\mathbf{a}} + \mathbf{a}\mathbf{l} + \mathbf{l}^{\mathrm{T}} \mathbf{a} \right) \mathbf{g}_j \end{aligned} \qquad \begin{aligned} \dot{a}^{ij} &= \dot{\overline{\mathbf{g}^i \mathbf{a} \mathbf{g}^j}} \\ &= -\mathbf{l}^{\mathrm{T}} \mathbf{g}^i \mathbf{a} \mathbf{g}^j + \mathbf{g}^i \dot{\mathbf{a}} \mathbf{g}^j - \mathbf{g}^i \mathbf{a} \mathbf{l}^{\mathrm{T}} \mathbf{g}^j \\ &= \mathbf{g}^i \left(\dot{\mathbf{a}} - \mathbf{l}\mathbf{a} - \mathbf{a}\mathbf{l}^{\mathrm{T}} \right) \mathbf{g}^j \end{aligned} \tag{2.12.13}$$

The convected derivatives are thus

$$\begin{aligned} \dot{a}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j &= \dot{\mathbf{a}} + \mathbf{a}\mathbf{l} + \mathbf{l}^{\mathrm{T}} \mathbf{a} \\ \dot{a}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j &= \dot{\mathbf{a}} - \mathbf{l}\mathbf{a} - \mathbf{a}\mathbf{l}^{\mathrm{T}} \end{aligned} \tag{2.12.14}$$

As will be seen below, these quantities are **Lie derivatives** of the tensor **a**.

The time derivative of the components can be expressed in an alternative way, by expressing the spatial base vectors $\mathbf{g}_i, \mathbf{g}^i$ in terms of the material base vectors $\mathbf{G}_i, \mathbf{G}^i$; using Eqns. 2.10.23:

$$\begin{aligned} \dot{a}_{ij} &= \dot{\overline{\mathbf{g}_i \mathbf{a} \mathbf{g}_j}} \\ &= \dot{\overline{\mathbf{F}\mathbf{G}_i \mathbf{a} \mathbf{F}\mathbf{G}_i}} \quad , \\ &= \mathbf{G}_i \, \dot{\overline{\mathbf{F}^{\mathrm{T}} \mathbf{a} \mathbf{F}}} \, \mathbf{G}_i \end{aligned} \qquad \begin{aligned} \dot{a}^{ij} &= \dot{\overline{\mathbf{g}^i \mathbf{a} \mathbf{g}^j}} \\ &= \dot{\overline{\mathbf{F}^{-\mathrm{T}} \mathbf{G}^i \mathbf{a} \mathbf{F}^{-\mathrm{T}} \mathbf{G}^j}} \\ &= \mathbf{G}^i \, \dot{\overline{\mathbf{F}^{-1} \mathbf{a} \mathbf{F}^{-\mathrm{T}}}} \, \mathbf{G}^j \end{aligned} \tag{2.12.15}$$

So, as an alternative to Eqns. 2.12.14,

$$\begin{aligned} \dot{a}_{ij} \mathbf{G}^i \otimes \mathbf{G}^j &= \dot{\overline{\mathbf{F}^{\mathrm{T}} \mathbf{a} \mathbf{F}}} \\ \dot{a}^{ij} \mathbf{G}_i \otimes \mathbf{G}_j &= \dot{\overline{\mathbf{F}^{-1} \mathbf{a} \mathbf{F}^{-\mathrm{T}}}} \end{aligned} \tag{2.12.16}$$

As will be seen next, the quantities on the right are the material time derivatives of the **pull-back** of the tensor **a**.

Example

Considering again Example 1 which was worked through in detail in §2.10, suppose we have a shearing deformation as shown in Fig. 2.12.1 (this is Fig. 2.10.3).

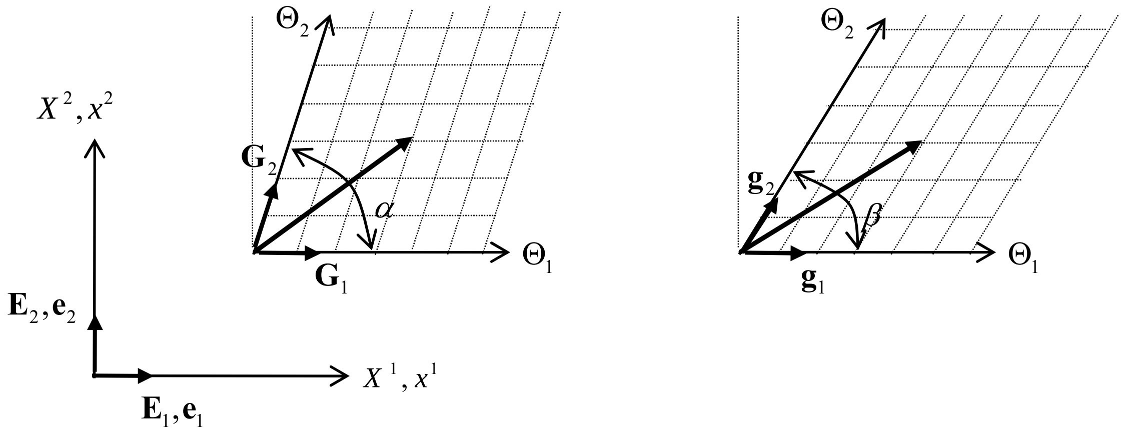


Figure 2.12.1: A pure shear deformation of one parallelogram into another

Let the shear angle β in Fig. 2.12.1 evolve over time according to

$$\beta = \alpha + \gamma t \tag{2.12.17}$$

From Eqns. 2.10.7, 2.10.11, the rates of change of the base vectors are

$$\begin{aligned} &\frac{d}{dt}\mathbf{g}_1 = \frac{d}{dt}\mathbf{e}_1 = 0, && \frac{d}{dt}\mathbf{g}_2 = \frac{d}{dt}\left(\frac{1}{\tan(\alpha+\gamma t)}\mathbf{e}_1 + \mathbf{e}_2\right) = -\frac{\gamma}{\sin^2(\alpha+\gamma t)}\mathbf{e}_1 \\ &\frac{d}{dt}\mathbf{g}^1 = \frac{d}{dt}\left(\mathbf{e}_1 - \frac{1}{\tan(\alpha+\gamma t)}\mathbf{e}_2\right) = +\frac{\gamma}{\sin^2(\alpha+\gamma t)}\mathbf{e}_2, && \frac{d}{dt}\mathbf{g}^2 = \frac{d}{dt}\mathbf{e}_2 = 0 \end{aligned} \tag{2.12.18}$$

The velocity gradient is, from Eqn. 2.11.5,

$$\begin{aligned} \mathbf{l} &= \dot{\mathbf{g}}_1 \otimes \mathbf{g}^1 + \dot{\mathbf{g}}_2 \otimes \mathbf{g}^2 \\ &= -\frac{\gamma}{\sin^2\beta}\mathbf{e}_1 \otimes \mathbf{e}_2 \\ &= \dot{\Pi}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (\mathbf{e}_i) \end{aligned} \tag{2.12.19}$$

where Π is given by Eqn. 2.10.26, and

$$\dot{\Pi}(t) = \frac{d}{dt}\left(\frac{1}{\tan(\alpha+\gamma t)} - \frac{1}{\tan(\alpha)}\right) = -\frac{\gamma}{\sin^2(\alpha+\gamma t)} \tag{2.12.20}$$

Considering again the vector $\mathbf{V}$ of Eqn. 2.10.44, $\mathbf{V} = \begin{bmatrix} V_x & V_y \end{bmatrix}^{\mathrm{T}} \left(\mathbf{E}_i\right)$, and its corresponding deformed vector $\mathbf{v}$ of Eqn. 2.10.47, $\mathbf{v} = \begin{bmatrix} V_x + \Pi V_y & V_y \end{bmatrix} \left(\mathbf{e}_i\right)$,

$$\dot{\mathbf{v}} = \dot{\Pi} \begin{bmatrix} V_y \\ 0 \end{bmatrix} \quad \left(\mathbf{e}_i\right), \tag{2.12.21}$$

The contravariant and covariant components of $\dot{\mathbf{v}}$ are

$$\dot{\mathbf{v}} = \hat{v}^i \mathbf{g}_i, \quad \hat{v}^i = \dot{\Pi} \begin{bmatrix} V_y \\ 0 \end{bmatrix}, \qquad \dot{\mathbf{v}} = \hat{v}_i \mathbf{g}^i, \quad \hat{v}_i = \dot{\Pi} \begin{bmatrix} V_y \\ \dfrac{1}{\tan\beta} V_y \end{bmatrix} \tag{2.12.22}$$

The "hat" on the $\hat{v}$ is to emphasise that (see Eqns. 2.12.5)

$$\hat{v}^i = \dot{\mathbf{v}} \cdot \mathbf{g}^i \neq \dot{v}^i = \dot{\overline{\mathbf{v} \cdot \mathbf{g}^i}}, \qquad \hat{v}_i = \dot{\mathbf{v}} \cdot \mathbf{g}_i \neq \dot{v}_i = \dot{\overline{\mathbf{v} \cdot \mathbf{g}_i}} \tag{2.12.23}$$

From Eqns. 2.12.6, the convected derivatives are

$$\begin{aligned} \dot{\mathbf{v}} - \mathbf{l}\mathbf{v} &= \dot{\Pi} \left\{ \begin{bmatrix} V_y \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_x + \Pi V_y \\ V_y \end{bmatrix} \right\}, & \dot{\mathbf{v}} + \mathbf{l}^{\mathrm{T}}\mathbf{v} &= \dot{\Pi} \left\{ \begin{bmatrix} V_y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} V_x + \Pi V_y \\ V_y \end{bmatrix} \right\} \\ &= \dot{\Pi} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & &= \dot{\Pi} \begin{bmatrix} V_y \\ V_x + \Pi V_y \end{bmatrix} \end{aligned} \tag{2.12.23}$$

Thus $\dot{\mathbf{v}} - \mathbf{l}\mathbf{v} = 0$, which, from Eqn. 2.12.6, implies that $\dot{v}^i = 0$. This is the expected result: the contravariant components do not change over time. They are always $\begin{bmatrix} V_x - V_y / \tan\alpha & V_y \end{bmatrix}$, as given by Eqn. 2.10.47b.

Consider now an example tensor

$$\mathbf{A} = \begin{bmatrix} A_{xx} & A_{xy} \\ A_{yx} & A_{yy} \end{bmatrix} \quad \left(\mathbf{E}_i\right) \tag{2.12.24}$$

The covariant and contravariant components are

$$\mathbf{A} = \begin{bmatrix} A_{xx} - \frac{1}{\tan\alpha}\left(A_{xy} + A_{yx}\right) + \frac{1}{\tan^2\alpha} A_{yy} & A_{xy} - \frac{1}{\tan\alpha} A_{yy} \\ A_{yx} - \frac{1}{\tan\alpha} A_{yy} & A_{yy} \end{bmatrix} \quad \left(\mathbf{G}_i\right)$$

$$\mathbf{A} = \begin{bmatrix} A_{xx} & A_{xy} + A_{xx}\frac{1}{\tan\alpha} \\ A_{yx} + A_{xx}\frac{1}{\tan\alpha} & A_{yy} + \frac{1}{\tan\alpha}\left(A_{xy} + A_{yx}\right) + \frac{1}{\tan^2\alpha} A_{xx} \end{bmatrix} \quad \left(\mathbf{G}^i\right) \tag{2.12.25}$$

This deforms to (with **F** given by Eqn. 2.10.25)

$$\mathbf{a} = \begin{bmatrix} 1 & \Pi \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_{xx} & A_{xy} \\ A_{yx} & A_{yy} \end{bmatrix} = \begin{bmatrix} A_{xx} + \Pi A_{yx} & A_{xy} + \Pi A_{yy} \\ A_{yx} & A_{yy} \end{bmatrix} \quad \left(\mathbf{e}_i\right) \tag{2.12.25}$$

Now

$$\begin{aligned} \mathbf{FA} &= \left(\mathbf{g}_i \otimes \mathbf{G}^i\right) A^{mn} \mathbf{G}_m \otimes \mathbf{G}_n \\ &= A^{ij} \mathbf{g}_i \otimes \mathbf{G}_j \end{aligned} \tag{2.12.26}$$

Converting between the various convected base vectors using Eqns. 2.10.7-8, 2.10.11-12, the contravariant and covariant and components are $\mathbf{a} = a^{ij}\mathbf{g}_i \otimes \mathbf{g}_j$, $\mathbf{a} = a_{ij}\mathbf{g}^i \otimes \mathbf{g}^j$:

$$a^{ij} = \begin{bmatrix} A_{xx} - \frac{1}{\tan\beta} A_{xy} - \frac{1}{\tan\alpha} A_{yx} + \frac{1}{\tan\alpha}\frac{1}{\tan\beta} A_{yy} & A_{xy} - A_{yy}\frac{1}{\tan\alpha} \\ A_{yx} - \frac{1}{\tan\beta} A_{yy} & A_{yy} \end{bmatrix}$$

$$a_{ij} = \begin{bmatrix} A_{xx} + A_{yx}\Pi & A_{xy} + A_{yy}\Pi + \frac{1}{\tan\beta}\left(A_{xx} + A_{yx}\Pi\right) \\ A_{yx} + \frac{1}{\tan\beta}\left(A_{xx} + A_{yx}\Pi\right) & A_{yy} + \frac{1}{\tan\beta} A_{yx} + \frac{1}{\tan\beta}\left(A_{xy} + A_{yy}\Pi\right) + \frac{1}{\tan^2\beta}\left(A_{xx} + A_{yx}\Pi\right) \end{bmatrix} \tag{2.12.27}$$

Also,

$$\dot{\mathbf{a}} = \dot{\Pi} \begin{bmatrix} A_{yx} & A_{yy} \\ 0 & 0 \end{bmatrix} \quad \left(\mathbf{e}_i\right), \tag{2.12.28}$$

and the contravariant and covariant components are

$$\dot{\mathbf{a}} = \hat{a}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j, \quad \hat{a}^{ij} = \dot{\Pi}\begin{bmatrix} A_{yx} - \dfrac{1}{\tan\beta}A_{yy} & A_{yy} \\ 0 & 0 \end{bmatrix}$$

$$\dot{\mathbf{a}} = \hat{a}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j, \quad \hat{a}_{ij} = \dot{\Pi}\begin{bmatrix} A_{yx} & \dfrac{1}{\tan\beta}A_{yx} + A_{yy} \\ \dfrac{1}{\tan\beta}A_{yx} & \dfrac{1}{\tan\beta}\left(\dfrac{1}{\tan\beta}A_{yx} + A_{yy}\right) \end{bmatrix} \tag{2.12.29}$$

Again, the "hat" emphasises that (see Eqns. 2.12.13)

$$\hat{a}^{ij} = \mathbf{g}^i\dot{\mathbf{a}}\mathbf{g}^j \neq \dot{a}^{ij} = \overline{\mathbf{g}^i\mathbf{a}\mathbf{g}^j}^{\boldsymbol{\cdot}}, \qquad \hat{a}_{ij} = \mathbf{g}_i\dot{\mathbf{a}}\mathbf{g}_j \neq \dot{a}_{ij} = \overline{\mathbf{g}_i\mathbf{a}\mathbf{g}_j}^{\boldsymbol{\cdot}} \tag{2.12.30}$$

Now

$$\dot{\mathbf{a}} - \mathbf{l}\mathbf{a} - \mathbf{a}\mathbf{l}^{\mathrm{T}} = \dot{\Pi}\begin{bmatrix} -A_{xy} - A_{yy}\Pi & 0 \\ -A_{yy} & 0 \end{bmatrix}$$

$$\dot{\mathbf{a}} + \mathbf{a}\mathbf{l} + \mathbf{l}^{\mathrm{T}}\mathbf{a} = \dot{\Pi}\begin{bmatrix} A_{yx} & A_{xx} + A_{yx}\Pi + A_{yy} \\ A_{xx} + A_{yx}\Pi & A_{xy} + A_{yx} + A_{yy}\Pi \end{bmatrix} \tag{2.12.31}$$

Thus $\dot{\mathbf{a}} - \mathbf{l}\mathbf{a} - \mathbf{a}\mathbf{l}^{\mathrm{T}} = 0$, i.e. $\dot{a}^{ij} = 0$, only when $A_{xy} = A_{yy} = 0$, which is consistent with Eqn. 2.12.27a (only constant terms, independent of β remain in that case).

2.12.2 Push-Forward and Pull-Back

Next are defined the push-forward and pull-back of vectors and tensors, which will lead into the concept of Lie derivatives, which relate back to what was just discussed above regarding convected derivatives.

Vectors

Consider a vector **V** given in terms of the reference configuration base vectors:

$$\begin{aligned} \mathbf{V} &= V_i\left(\Theta^j\right)\mathbf{G}^i \\ &= V^i\left(\Theta^j\right)\mathbf{G}_i \end{aligned} \tag{2.12.32}$$

The **push-forward**, symbolised by $\chi_*(\bullet)$, is defined to be the vector with *the same components*, but with respect to the current configuration base vectors. There are 2 push-forward operations, depending on the type of components used; the symbol b is used for covariant components V_i and the symbol # for contravariant components V^i; using 2.10.23,

$$\boxed{\begin{aligned}\chi_*(\mathbf{V})^b &\equiv V_i\mathbf{g}^i = V_i\mathbf{F}^{-\mathrm{T}}\mathbf{G}^i = \mathbf{F}^{-\mathrm{T}}\mathbf{V}\\ \chi_*(\mathbf{V})^\# &\equiv V^i\mathbf{g}_i = V^i\mathbf{F}\mathbf{G}_i = \mathbf{F}\mathbf{V}\end{aligned}}$$ **Push-forward of Vector** (2.12.33)

Eqn. 2.12.33b says that the push forward of the contravariant form of **V** is simply **FV**. In other words, the push forward here is the actual corresponding vector in the deformed configuration, $\mathbf{v} = \mathbf{F}\mathbf{V} = v^i\left(\Theta^j\right)\mathbf{g}_i$, and, as a consequence of the definitions, $V^i = v^i$, as illustrated in Fig. 2.12.2.

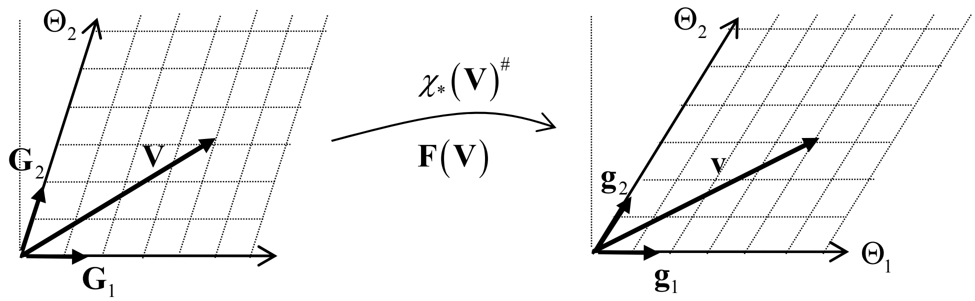


Figure 2.12.2: The push-forward of a vector V

A special case of Eqn. 2.12.33b is the push forward of a line element in the reference configuration, giving the corresponding line element in the current configuration:

$$\chi_*(d\mathbf{X})^\# = d\Theta^i\mathbf{g}_i = d\mathbf{x}. \qquad (2.12.34)$$

Similarly, consider a vector **v** given in terms of the current configuration basis:

$$\mathbf{v} = v_i\mathbf{g}^i = v^i\mathbf{g}_i \qquad (2.12.35)$$

The **pull-back** of **v**, $\chi_*^{-1}(\mathbf{v})$, is defined to be the vector with components v_i (or v^i) with respect to the reference configuration base vectors $\mathbf{G}^i$ (or $\mathbf{G}_i$). Using 2.10.23,

$$\boxed{\begin{aligned}\chi_*^{-1}(\mathbf{v})^b &= v_i\mathbf{G}^i = v_i\mathbf{F}^{\mathrm{T}}\mathbf{g}^i = \mathbf{F}^{\mathrm{T}}\mathbf{v}\\ \chi_*^{-1}(\mathbf{v})^\# &= v^i\mathbf{G}_i = v^i\mathbf{F}^{-1}\mathbf{g}_i = \mathbf{F}^{-1}\mathbf{v}\end{aligned}}$$ **Pull-back of a vector** (2.12.36)

and, for a line element in the current configuration,

$$\chi_*^{-1}(d\mathbf{x})^\# = dx^i\mathbf{G}_i = \mathbf{F}^{-1}d\mathbf{x} = d\mathbf{X}. \qquad (2.12.37)$$

Note that a push-forward and pull-back applied successively to a vector with the same component type will result in the initial vector.

From the above, for two material vectors **U** and **V** and two spatial vectors **u** and **v**,

$$\begin{aligned}\mathbf{U}\cdot\mathbf{V} &= \chi_*(\mathbf{U})^b \cdot \chi_*(\mathbf{V})^\# = \chi_*(\mathbf{U})^\# \cdot \chi_*(\mathbf{V})^b \\ \mathbf{u}\cdot\mathbf{v} &= \chi_*^{-1}(\mathbf{u})^b \cdot \chi_*^{-1}(\mathbf{v})^\# = \chi_*^{-1}(\mathbf{u})^\# \cdot \chi_*^{-1}(\mathbf{v})^b\end{aligned} \tag{2.12.38}$$

For example, as a special case of this, in the reference configuration, $\mathbf{G}_1$ and $\mathbf{G}^2$ are perpendicular: $\mathbf{G}_1 \cdot \mathbf{G}^2 = 0$. Pushing forward these vectors, we get from Eqn. 2.12.33: $\mathbf{F}\mathbf{G}_1 = \mathbf{g}_1$ and $\mathbf{F}^{-\mathrm{T}}\mathbf{G}^2 = \mathbf{g}^2$, and again $\chi_*(\mathbf{G}_1)^\# \cdot \chi_*(\mathbf{G}^2)^b = \mathbf{g}_1 \cdot \mathbf{g}^2 = 0$.

Tensors

Consider a material tensor **A**:

$$\mathbf{A} = A_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = A^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = A^i_{\cdot j}\mathbf{G}_i \otimes \mathbf{G}^j = A_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j \tag{2.12.39}$$

As for the vector, the push-forward of **A**, $\chi_*(\mathbf{A})$, is defined to be the tensor with the same components, but with respect to the deformed base vectors. Thus, using 2.10.23,

$$\boxed{\begin{aligned}\chi_*(\mathbf{A})^b &= A_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = A_{ij}\left(\mathbf{F}^{-\mathrm{T}}\mathbf{G}^i \otimes \mathbf{F}^{-\mathrm{T}}\mathbf{G}^j\right) = \mathbf{F}^{-\mathrm{T}}\mathbf{A}\mathbf{F}^{-1} \\ \chi_*(\mathbf{A})^\# &= A^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = A^{ij}\left(\mathbf{F}\mathbf{G}_i \otimes \mathbf{F}\mathbf{G}_j\right) = \mathbf{F}\mathbf{A}\mathbf{F}^{\mathrm{T}} \\ \chi_*(\mathbf{A})^\backslash &= A^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j = A^i_{\cdot j}\left(\mathbf{F}\mathbf{G}_i \otimes \mathbf{F}^{-\mathrm{T}}\mathbf{G}^j\right) = \mathbf{F}\mathbf{A}\mathbf{F}^{-1} \\ \chi_*(\mathbf{A})^/ &= A_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = A_i^{\cdot j}\left(\mathbf{F}^{-\mathrm{T}}\mathbf{G}^i \otimes \mathbf{F}\mathbf{G}_j\right) = \mathbf{F}^{-\mathrm{T}}\mathbf{A}\mathbf{F}^{\mathrm{T}}\end{aligned}} \quad \textbf{Push-forward of Tensor} \tag{2.12.40}$$

Similarly, consider a spatial tensor **a**:

$$\mathbf{a} = a_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = a^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = a^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j = a_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j \tag{2.12.41}$$

The pull-back is

$$\boxed{\begin{aligned}
\chi_*^{-1}(\mathbf{a})^{b} &= a_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = a_{ij}\left(\mathbf{F}^{\mathrm{T}}\mathbf{g}^i \otimes \mathbf{F}^{\mathrm{T}}\mathbf{g}^j\right) &&= \mathbf{F}^{\mathrm{T}}\mathbf{a}\mathbf{F} \\
\chi_*^{-1}(\mathbf{a})^{\#} &= a^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = a^{ij}\left(\mathbf{F}^{-1}\mathbf{g}_i \otimes \mathbf{F}^{-1}\mathbf{g}_j\right) &&= \mathbf{F}^{-1}\mathbf{a}\mathbf{F}^{-\mathrm{T}} \\
\chi_*^{-1}(\mathbf{a})^{\backslash} &= a^{i}_{\cdot j}\mathbf{G}_i \otimes \mathbf{G}^j = a^{i}_{\cdot j}\left(\mathbf{F}^{-1}\mathbf{g}_i \otimes \mathbf{F}^{\mathrm{T}}\mathbf{g}^j\right) &&= \mathbf{F}^{-1}\mathbf{a}\mathbf{F} \\
\chi_*^{-1}(\mathbf{a})^{/} &= a_{i}^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = a_{i}^{\cdot j}\left(\mathbf{F}^{\mathrm{T}}\mathbf{g}^i \otimes \mathbf{F}^{-1}\mathbf{g}_j\right) &&= \mathbf{F}^{\mathrm{T}}\mathbf{a}\mathbf{F}^{-\mathrm{T}}
\end{aligned}} \qquad \textbf{Pull-back of Tensor} \qquad (2.12.42)$$

The first of these, $\mathbf{F}^{\mathrm{T}}\mathbf{a}\mathbf{F}$, is called the **covariant pull-back**, whereas the second, $\mathbf{F}^{-1}\mathbf{a}\mathbf{F}^{-\mathrm{T}}$, is called the **contravariant pull-back**.

Since $\mathbf{F}$ maps material vectors to spatial vectors, $\mathbf{a}$ maps spatial vectors to spatial vectors, and $\mathbf{F}^{\mathrm{T}}$ maps spatial vectors to material vectors, it follows that the pull-back $\mathbf{F}^{\mathrm{T}}\mathbf{a}\mathbf{F}$ maps material vectors to material vectors, and so is a material tensor field, and similarly for the other three pull-backs.

Time Derivatives

It will be recognised that the expressions for the pull backs of a spatial covariant tensor and spatial contravariant tensor in Eqns. 2.12.42a,b are those appearing in Eqns. 2.12.16. Keeping in mind Eqn. 2.12.14, one sees that, for a spatial tensor in terms of covariant components, $\mathbf{a} = a_{ij}\mathbf{g}^i \otimes \mathbf{g}^j$, and contravariant components, $\mathbf{a} = a^{ij}\mathbf{g}_i \otimes \mathbf{g}_j$,

$$\begin{aligned}
\dot{a}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j &= \left(\dot{\overline{\mathbf{F}^{\mathrm{T}}\mathbf{a}\mathbf{F}}}\right)\mathbf{g}^i \otimes \mathbf{g}^j &&= \dot{\mathbf{a}} + \mathbf{a}\mathbf{l} + \mathbf{l}^{\mathrm{T}}\mathbf{a} \\
\dot{a}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j &= \left(\dot{\overline{\mathbf{F}^{-1}\mathbf{a}\mathbf{F}^{-\mathrm{T}}}}\right)\mathbf{g}_i \otimes \mathbf{g}_j &&= \dot{\mathbf{a}} - \mathbf{l}\mathbf{a} - \mathbf{a}\mathbf{l}^{\mathrm{T}}
\end{aligned} \qquad (2.12.43)$$

Other Push-Forward and Pull-Back relations for Vectors and Tensors

Here follow some relations involving the push-forward and pull-backs of tensors.

For two material tensors $\mathbf{A}$ and $\mathbf{B}$ and two spatial tensors $\mathbf{a}$ and $\mathbf{b}$, the scalar product is

$$\begin{aligned}
\mathbf{A} : \mathbf{B} &= A_{ij}B^{ij} = A^{ij}B_{ij} = A^{i}_{\cdot j}B_{i}^{\cdot j} = A_{i}^{\cdot j}B^{i}_{\cdot j} \\
\mathbf{a} : \mathbf{b} &= a_{ij}b^{ij} = a^{ij}b_{ij} = a^{i}_{\cdot j}b_{i}^{\cdot j} = a_{i}^{\cdot j}b^{i}_{\cdot j}
\end{aligned} \qquad (2.12.44)$$

This scalar product then push-forwards and pull-backs as {▲Problem 1}

$$
\begin{aligned}
\mathbf{A}:\mathbf{B} &= \chi_*(\mathbf{A})^\flat : \chi_*(\mathbf{B})^\# = \chi_*(\mathbf{A})^\# : \chi_*(\mathbf{B})^\flat \\
&= \chi_*(\mathbf{A})^{/} : \chi_*(\mathbf{B})^{\backslash} = \chi_*(\mathbf{A})^{\backslash} : \chi_*(\mathbf{B})^{/} \\
\mathbf{a}:\mathbf{b} &= \chi_*^{-1}(\mathbf{a})^\flat : \chi_*^{-1}(\mathbf{b})^\# = \chi_*^{-1}(\mathbf{a})^\# : \chi_*^{-1}(\mathbf{b})^\flat \\
&= \chi_*^{-1}(\mathbf{a})^{/} : \chi_*^{-1}(\mathbf{b})^{\backslash} = \chi_*^{-1}(\mathbf{a})^{\backslash} : \chi_*^{-1}(\mathbf{b})^{/}
\end{aligned}
\tag{2.12.45}
$$

For material tensor $\mathbf{A}$ and material vectors $\mathbf{U},\mathbf{V}$, and spatial tensor $\mathbf{a}$ and spatial vectors $\mathbf{u},\mathbf{v}$,

$$
\begin{aligned}
\mathbf{UAV} &= U_i A^{ij} V_j = U^i A_{ij} V^j = U_i A^{i}_{\cdot j} V^j = U^i A_i^{\cdot j} V_j \\
\mathbf{uav} &= u_i a^{ij} v_j = u^i a_{ij} v^j = u_i a^{i}_{\cdot j} v^j = u^i a_i^{\cdot j} v_j
\end{aligned}
\tag{2.12.46}
$$

Then

$$
\begin{aligned}
\mathbf{UAV} &= \chi_*(\mathbf{U})^\flat \chi_*(\mathbf{A})^\# \chi_*(\mathbf{V})^\flat = \chi_*(\mathbf{U})^\# \chi_*(\mathbf{A})^\flat \chi_*(\mathbf{V})^\# \\
&= \chi_*(\mathbf{U})^\flat \chi_*(\mathbf{A})^{\backslash} \chi_*(\mathbf{V})^\# = \chi_*(\mathbf{U})^\# \chi_*(\mathbf{A})^{/} \chi_*(\mathbf{V})^\flat \\
\mathbf{uav} &= \chi_*^{-1}(\mathbf{u})^\flat \chi_*^{-1}(\mathbf{a})^\# \chi_*^{-1}(\mathbf{v})^\flat = \chi_*^{-1}(\mathbf{u})^\# \chi_*^{-1}(\mathbf{a})^\flat \chi_*^{-1}(\mathbf{v})^\# \\
&= \chi_*^{-1}(\mathbf{u})^\flat \chi_*^{-1}(\mathbf{a})^{\backslash} \chi_*^{-1}(\mathbf{v})^\# = \chi_*^{-1}(\mathbf{u})^\# \chi_*^{-1}(\mathbf{a})^{/} \chi_*^{-1}(\mathbf{v})^\flat
\end{aligned}
\tag{2.12.47}
$$

For material tensor **A** and material vector **V**, and spatial tensor **a** and spatial vector **v**, the contractions $\mathbf{AV}$ and $\mathbf{av}$ are

$$
\begin{aligned}
\mathbf{AV} &= A_{ij} V^j = A_i^{\cdot j} V_j = A^{i}_{\cdot j} V^j = A^{ij} V_j \\
\mathbf{av} &= a_{ij} v^j = a_i^{\cdot j} v_j = a^{i}_{\cdot j} v^j = a^{ij} v_j
\end{aligned}
\tag{2.12.48}
$$

and so transform as

$$
\begin{aligned}
\chi_*(\mathbf{AV})^\flat &= \chi_*(\mathbf{A})^\flat \chi_*(\mathbf{V})^\# = \chi_*(\mathbf{A})^{/} \chi_*(\mathbf{V})^\flat \\
\chi_*(\mathbf{AV})^\# &= \chi_*(\mathbf{A})^\# \chi_*(\mathbf{V})^\flat = \chi_*(\mathbf{A})^{\backslash} \chi_*(\mathbf{V})^\# \\
\\
\chi_*^{-1}(\mathbf{av})^\flat &= \chi_*^{-1}(\mathbf{a})^\flat \chi_*^{-1}(\mathbf{v})^\# = \chi_*^{-1}(\mathbf{a})^{/} \chi_*^{-1}(\mathbf{v})^\flat \\
\chi_*^{-1}(\mathbf{av})^\# &= \chi_*^{-1}(\mathbf{a})^\# \chi_*^{-1}(\mathbf{v})^\flat = \chi_*^{-1}(\mathbf{a})^{\backslash} \chi_*^{-1}(\mathbf{v})^\#
\end{aligned}
\tag{2.12.49}
$$

Finally, for material tensors **A**, **B** and spatial tensors **a**, **b**,

$$
\begin{aligned}
\mathbf{AB} &= A_{ik} B^{kj} \mathbf{G}^i \otimes \mathbf{G}_j = A_i^{\cdot k} B_k^{\cdot j} \mathbf{G}^i \otimes \mathbf{G}_j = A_i^{\cdot k} B_{kj} \mathbf{G}^i \otimes \mathbf{G}^j = A_{ik} B^{k}_{\cdot j} \mathbf{G}^i \otimes \mathbf{G}^j = \cdots \\
\mathbf{ab} &= a_{ik} b^{kj} \mathbf{g}^i \otimes \mathbf{g}_j = a_i^{\cdot k} b_k^{\cdot j} \mathbf{g}^i \otimes \mathbf{g}_j = a_i^{\cdot k} b_{kj} \mathbf{g}^i \otimes \mathbf{g}^j = a_{ik} b^{k}_{\cdot j} \mathbf{g}^i \otimes \mathbf{g}^j = \cdots
\end{aligned}
\tag{2.12.50}
$$

and so

$$\begin{aligned}
&\chi_*(\mathbf{AB})^{/} = \chi_*(\mathbf{A})^{b}\,\chi_*(\mathbf{B})^{\#} = \chi_*(\mathbf{A})^{/}\,\chi_*(\mathbf{B})^{/} \\
&\chi_*(\mathbf{AB})^{b} = \chi_*(\mathbf{A})^{/}\,\chi_*(\mathbf{B})^{b} = \chi_*(\mathbf{A})^{b}\,\chi_*(\mathbf{B})^{\backslash} \\
&\vdots \\
&\chi_*^{-1}(\mathbf{ab})^{/} = \chi_*^{-1}(\mathbf{a})^{b}\,\chi_*^{-1}(\mathbf{b})^{\#} = \chi_*^{-1}(\mathbf{a})^{/}\,\chi_*^{-1}(\mathbf{b})^{/} \\
&\chi_*^{-1}(\mathbf{ab})^{b} = \chi_*^{-1}(\mathbf{a})^{b}\,\chi_*^{-1}(\mathbf{b})^{\#} = \chi_*^{-1}(\mathbf{a})^{/}\,\chi_*^{-1}(\mathbf{b})^{/}
\end{aligned} \tag{2.12.51}$$

Push-Forward and Pull-Back operations for Strain Tensors

The push-forward of the covariant right Cauchy-Green strain and its contravariant inverse are

$$\begin{aligned}
\chi_*(\mathbf{C})^{b} &= C_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = \mathbf{F}^{-\mathrm{T}}\mathbf{C}\mathbf{F}^{-1} \\
\chi_*\left(\mathbf{C}^{-1}\right)^{\#} &= \left(C^{-1}\right)^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = \mathbf{F}\mathbf{C}\mathbf{F}^{\mathrm{T}}
\end{aligned}. \tag{2.12.52}$$

From 2.10.39, $C_{ij} = g_{ij}$, the covariant components of the identity tensor expressed in terms of the convected base vectors in the current configuration, i.e. the spatial metric tensor, $\mathbf{g} = g_{ij}\mathbf{g}^i \otimes \mathbf{g}^i$, and $\left(C^{-1}\right)^{ij} = g^{ij}$, the contravariant components of **g**. Thus the push-forward of covariant **C** is **g** and the pull-back of covariant **g** is **C**, and the push-forward of contravariant $\mathbf{C}^{-1}$ is **g** and the pull-back of contravariant **g** is $\mathbf{C}^{-1}$:

$$\boxed{\begin{aligned}
\chi_*(\mathbf{C})^{b} = \mathbf{g}, \quad \chi_*^{-1}(\mathbf{g})^{b} = \mathbf{C} \\
\chi_*\left(\mathbf{C}^{-1}\right)^{\#} = \mathbf{g}, \quad \chi_*^{-1}(\mathbf{g})^{\#} = \mathbf{C}^{-1}
\end{aligned}} \tag{2.12.53}$$

Push-forward of the right Cauchy-Green strain

Similarly, the pull-back of covariant $\mathbf{b}^{-1}$ is **G** and the push-forward of covariant **G** is $\mathbf{b}^{-1}$**,** and the pull-back of contravariant **b** is **G** and the push-forward of contravariant **G** is **b**.

$$\boxed{\begin{aligned}
\chi_*(\mathbf{G})^{b} = \mathbf{b}^{-1}, \quad \chi_*^{-1}\left(\mathbf{b}^{-1}\right)^{b} = \mathbf{G} \\
\chi_*(\mathbf{G})^{\#} = \mathbf{b}, \qquad \chi_*^{-1}(\mathbf{b})^{\#} = \mathbf{G}
\end{aligned}} \tag{2.12.54}$$

Pull-back of the left Cauchy-Green strain

For the covariant form of the Green-Lagrange strain, the push-forward is

$$\chi_*(\mathbf{E})^{b} = E_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = \mathbf{F}^{-\mathrm{T}}\mathbf{E}\mathbf{F}^{-1}. \tag{2.12.55}$$

From 2.10.43, $E_{ij} = e_{ij}$, the covariant components of the Euler-Almansi strain tensor, and so the push-forward of covariant **E** is **e** and the pull-back of covariant **e** is **E**.

$$\boxed{\chi_*(\mathbf{E})^b = \mathbf{e}, \qquad \chi_*^{-1}(\mathbf{e})^b = \mathbf{E}}. \tag{2.12.56}$$

Push-forward of the Green-Lagrange strain
Pull-back of the Euler-Almansi strain

Push-Forward and Pull-Back with Polar Decomposition Intermediate Configurations

Pull backs and push-forwards can be defined relative to any two configurations. Consider the polar decomposition and the intermediate configurations discussed in §2.10 (see Fig. 2.10.11). Effectively, we are replacing $\mathbf{F}$ with $\mathbf{R}$: pushing forward a material tensor $\mathbf{A}$ from the reference configuration $\{\mathbf{G}_i\}$ to the configuration $\{\hat{\mathbf{G}}_i\}$ leads to

$$\begin{aligned}
\chi_*(\mathbf{A})^b{}_{\mathbf{R}(\mathbf{G})} &= A_{ij}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}^j = A_{ij}\left(\mathbf{R}^{-\mathrm{T}}\mathbf{G}^i \otimes \mathbf{R}^{-\mathrm{T}}\mathbf{G}^j\right) = \mathbf{R}^{-\mathrm{T}}\mathbf{A}\mathbf{R}^{-1} = \mathbf{R}\mathbf{A}\mathbf{R}^{\mathrm{T}} \\
\chi_*(\mathbf{A})^\#{}_{\mathbf{R}(\mathbf{G})} &= A^{ij}\hat{\mathbf{G}}_i \otimes \hat{\mathbf{G}}_j = A^{ij}\left(\mathbf{R}\mathbf{G}_i \otimes \mathbf{R}\mathbf{G}_j\right) = \mathbf{R}\mathbf{A}\mathbf{R}^{\mathrm{T}} \\
\chi_*(\mathbf{A})^\backslash{}_{\mathbf{R}(\mathbf{G})} &= A^i_{\cdot j}\hat{\mathbf{G}}_i \otimes \hat{\mathbf{G}}^j = A^i_{\cdot j}\left(\mathbf{R}\mathbf{G}_i \otimes \mathbf{R}^{-\mathrm{T}}\mathbf{G}^j\right) = \mathbf{R}\mathbf{A}\mathbf{R}^{-1} = \mathbf{R}\mathbf{A}\mathbf{R}^{\mathrm{T}} \\
\chi_*(\mathbf{A})^/{}_{\mathbf{R}(\mathbf{G})} &= A_i^{\cdot j}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}_j = A_i^{\cdot j}\left(\mathbf{R}^{-\mathrm{T}}\mathbf{G}^i \otimes \mathbf{R}\mathbf{G}_j\right) = \mathbf{R}^{-\mathrm{T}}\mathbf{A}\mathbf{R}^{\mathrm{T}} = \mathbf{R}\mathbf{A}\mathbf{R}^{\mathrm{T}}
\end{aligned} \; . \tag{2.12.57}$$

Note that the result is the same regardless of whether one is using the covariant, contravariant or mixed forms.

Similarly, the pull back of a tensor $\hat{\mathbf{A}}$ from the intermediate configuration $\{\hat{\mathbf{G}}_i\}$ to the reference configuration $\{\mathbf{G}_i\}$ is

$$\begin{aligned}
\chi_*^{-1}(\hat{\mathbf{A}})^b{}_{\mathbf{R}(\hat{\mathbf{G}})} &= \hat{A}_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = \mathbf{R}^{\mathrm{T}}\hat{\mathbf{A}}\mathbf{R} \\
\chi_*^{-1}(\hat{\mathbf{A}})^\#{}_{\mathbf{R}(\hat{\mathbf{G}})} &= \hat{A}^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = \mathbf{R}^{\mathrm{T}}\hat{\mathbf{A}}\mathbf{R} \\
\chi_*^{-1}(\hat{\mathbf{A}})^\backslash{}_{\mathbf{R}(\hat{\mathbf{G}})} &= \hat{A}^i_{\cdot j}\mathbf{G}_i \otimes \mathbf{G}^j = \mathbf{R}^{\mathrm{T}}\hat{\mathbf{A}}\mathbf{R} \\
\chi_*^{-1}(\hat{\mathbf{A}})^/{}_{\mathbf{R}(\hat{\mathbf{G}})} &= \hat{A}_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = \mathbf{R}^{\mathrm{T}}\hat{\mathbf{A}}\mathbf{R}
\end{aligned} \tag{2.12.58}$$

The push-forward of a tensor $\hat{\mathbf{a}}$ from $\{\hat{\mathbf{g}}_i\}$ to $\{\mathbf{g}\}$ and the corresponding pull-back of a spatial tensor $\mathbf{a}$ is

$$\begin{aligned}
\chi_*(\hat{\mathbf{a}})^b{}_{\mathbf{R}(\hat{\mathbf{g}})} &= \hat{a}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = \mathbf{R}\hat{\mathbf{a}}\mathbf{R}^{\mathrm{T}} & \chi_*^{-1}(\mathbf{a})^b{}_{\mathbf{R}(\mathbf{g})} &= a_{ij}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}^j = \mathbf{R}^{\mathrm{T}}\mathbf{a}\mathbf{R} \\
\chi_*(\hat{\mathbf{a}})^\#{}_{\mathbf{R}(\hat{\mathbf{g}})} &= \hat{a}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = \mathbf{R}\hat{\mathbf{a}}\mathbf{R}^{\mathrm{T}} & \chi_*^{-1}(\mathbf{a})^\#{}_{\mathbf{R}(\mathbf{g})} &= a^{ij}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}_j = \mathbf{R}^{\mathrm{T}}\mathbf{a}\mathbf{R} \\
\chi_*(\hat{\mathbf{a}})^\backslash{}_{\mathbf{R}(\hat{\mathbf{g}})} &= \hat{a}^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j = \mathbf{R}\hat{\mathbf{a}}\mathbf{R}^{\mathrm{T}}, & \chi_*^{-1}(\mathbf{a})^\backslash{}_{\mathbf{R}(\mathbf{g})} &= a^i_{\cdot j}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}^j = \mathbf{R}^{\mathrm{T}}\mathbf{a}\mathbf{R} \\
\chi_*(\hat{\mathbf{a}})^/{}_{\mathbf{R}(\hat{\mathbf{g}})} &= \hat{a}_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = \mathbf{R}\hat{\mathbf{a}}\mathbf{R}^{\mathrm{T}} & \chi_*^{-1}(\mathbf{a})^/{}_{\mathbf{R}(\mathbf{g})} &= a_i^{\cdot j}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}_j = \mathbf{R}^{\mathrm{T}}\mathbf{a}\mathbf{R}
\end{aligned} \tag{2.12.59}$$

The push-forwards and pull-backs due to the stretch tensors are

$$
\begin{aligned}
\chi_*(\mathbf{A})^b{}_{\mathbf{U}(\mathbf{G})} &= A_{ij}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}^j = A_{ij}\left(\mathbf{U}^{-\mathrm{T}}\mathbf{G}^i \otimes \mathbf{U}^{-\mathrm{T}}\mathbf{G}^j\right) = \mathbf{U}^{-\mathrm{T}}\mathbf{A}\mathbf{U}^{-1} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}^{-1} \\
\chi_*(\mathbf{A})^\#{}_{\mathbf{U}(\mathbf{G})} &= A^{ij}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}_j = A^{ij}\left(\mathbf{U}\mathbf{G}_i \otimes \mathbf{U}\mathbf{G}_j\right) \qquad = \mathbf{U}\mathbf{A}\mathbf{U}^{\mathrm{T}} = \mathbf{U}\mathbf{A}\mathbf{U} \\
\chi_*(\mathbf{A})^\backslash{}_{\mathbf{U}(\mathbf{G})} &= A^i_{\cdot j}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}^j = A^i_{\cdot j}\left(\mathbf{U}\mathbf{G}_i \otimes \mathbf{U}^{-\mathrm{T}}\mathbf{G}^j\right) \quad = \mathbf{U}\mathbf{A}\mathbf{U}^{-1} \\
\chi_*(\mathbf{A})^/{}_{\mathbf{U}(\mathbf{G})} &= A_i^{\cdot j}\hat{\mathbf{g}}^i \otimes \hat{\mathbf{g}}_j = A_i^{\cdot j}\left(\mathbf{U}^{-\mathrm{T}}\mathbf{G}^i \otimes \mathbf{U}\mathbf{G}_j\right) \quad = \mathbf{U}^{-\mathrm{T}}\mathbf{A}\mathbf{U}^{\mathrm{T}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}
\end{aligned}
\quad . \qquad (2.12.60)
$$

$$
\begin{aligned}
\chi_*^{-1}(\hat{\mathbf{a}})^b{}_{\mathbf{U}(\hat{\mathbf{g}})} &= \hat{a}_{ij}\mathbf{G}^i \otimes \mathbf{G}^j = \mathbf{U}\hat{\mathbf{a}}\mathbf{U} \\
\chi_*^{-1}(\hat{\mathbf{a}})^\#{}_{\mathbf{U}(\hat{\mathbf{g}})} &= \hat{a}^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = \mathbf{U}^{-1}\hat{\mathbf{a}}\mathbf{U}^{-1} \\
\chi_*^{-1}(\hat{\mathbf{a}})^\backslash{}_{\mathbf{U}(\hat{\mathbf{g}})} &= \hat{a}^i_{\cdot j}\mathbf{G}_i \otimes \mathbf{G}^j = \mathbf{U}^{-1}\hat{\mathbf{a}}\mathbf{U} \\
\chi_*^{-1}(\hat{\mathbf{a}})^/{}_{\mathbf{U}(\hat{\mathbf{g}})} &= \hat{a}_i^{\cdot j}\mathbf{G}^i \otimes \mathbf{G}_j = \mathbf{U}\hat{\mathbf{a}}\mathbf{U}^{-1}
\end{aligned}
\qquad (2.12.61)
$$

and

$$
\begin{aligned}
\chi_*(\hat{\mathbf{A}})^b{}_{\mathbf{v}(\hat{\mathbf{G}})} &= \hat{A}_{ij}\mathbf{g}^i \otimes \mathbf{g}^j = \mathbf{v}^{-1}\hat{\mathbf{A}}\mathbf{v}^{-1} \\
\chi_*(\hat{\mathbf{A}})^\#{}_{\mathbf{v}(\hat{\mathbf{G}})} &= \hat{A}^{ij}\mathbf{g}_i \otimes \mathbf{g}_j = \mathbf{v}\hat{\mathbf{A}}\mathbf{v} \\
\chi_*(\hat{\mathbf{A}})^\backslash{}_{\mathbf{v}(\hat{\mathbf{G}})} &= \hat{A}^i_{\cdot j}\mathbf{g}_i \otimes \mathbf{g}^j = \mathbf{v}\hat{\mathbf{A}}\mathbf{v}^{-1} \\
\chi_*(\hat{\mathbf{A}})^/{}_{\mathbf{v}(\hat{\mathbf{G}})} &= \hat{A}_i^{\cdot j}\mathbf{g}^i \otimes \mathbf{g}_j = \mathbf{v}^{-1}\hat{\mathbf{A}}\mathbf{v}
\end{aligned}
\quad , \quad
\begin{aligned}
\chi_*^{-1}(\mathbf{a})^b{}_{\mathbf{v}(\mathbf{g})} &= a_{ij}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}^j = \mathbf{v}\mathbf{a}\mathbf{v} \\
\chi_*^{-1}(\mathbf{a})^\#{}_{\mathbf{v}(\mathbf{g})} &= a^{ij}\hat{\mathbf{G}}_i \otimes \hat{\mathbf{G}}_j = \mathbf{v}^{-1}\mathbf{a}\mathbf{v}^{-1} \\
\chi_*^{-1}(\mathbf{a})^\backslash{}_{\mathbf{v}(\mathbf{g})} &= a^i_{\cdot j}\hat{\mathbf{G}}_i \otimes \hat{\mathbf{G}}^j = \mathbf{v}^{-1}\mathbf{a}\mathbf{v} \\
\chi_*^{-1}(\mathbf{a})^/{}_{\mathbf{v}(\mathbf{g})} &= a_i^{\cdot j}\hat{\mathbf{G}}^i \otimes \hat{\mathbf{G}}_j = \mathbf{v}\mathbf{a}\mathbf{v}^{-1}
\end{aligned}
\qquad (2.12.62)
$$

Push-forwards and pull-backs can also be defined using $\mathbf{F}^{\mathrm{T}}$ (in the place of $\mathbf{F}$) and these move between the intermediate configurations, $\hat{\mathbf{G}} \Leftrightarrow \hat{\mathbf{g}}$.

Recall Eqn. 2.10.64, which state that the covariant components of $\mathbf{U}, \mathbf{v}, \mathbf{U}^{-1}, \mathbf{v}^{-1}$ with respect to the bases $\mathbf{G}^i, \hat{\mathbf{G}}^i, \hat{\mathbf{g}}^i, \mathbf{g}^i$ respectively, are equal. This can be explained also in terms of push-forwards and pull-backs. For example, with $\mathbf{v} = \mathbf{R}\mathbf{U}\mathbf{R}^{\mathrm{T}}$ and $\mathbf{v}^{-1} = \mathbf{R}\mathbf{U}^{-1}\mathbf{R}^{\mathrm{T}}$, one can write (in fact these relations are valid for all component types)

$$
\mathbf{v} = \chi_*(\mathbf{U})_{\mathbf{R}(\mathbf{G})}, \qquad \mathbf{v}^{-1} = \chi_*\left(\mathbf{U}^{-1}\right)_{\mathbf{R}(\hat{\mathbf{g}})} \qquad (2.12.63)
$$

The first of these shows that the components of $\mathbf{U}$ with respect to $\mathbf{G}$ are the same as those of $\mathbf{v}$ with respect to $\hat{\mathbf{G}}$ (for all component types). The second shows that the components of $\mathbf{U}^{-1}$ with respect to $\hat{\mathbf{g}}$ are the same as those of $\mathbf{v}^{-1}$ with respect to $\mathbf{g}$.

As another example, with $\mathbf{C} = \mathbf{U}^2$,

$$\mathbf{C} = \chi_*^{-1}\left(\hat{\mathbf{g}}\right)^{b}\,\mathrm{U}(\hat{\mathbf{g}}), \qquad \mathbf{C}^{-1} = \chi_*^{-1}\left(\hat{\mathbf{g}}\right)^{\#}\,\mathrm{U}(\hat{\mathbf{g}}) \tag{2.12.64}$$

2.12.3 The Lie Time Derivative

The **Lie (time) derivative** is a concept of tensor analysis which is used to distinguish between the change in some quantity, and the change in that quantity excluding changes due to the motion/configuration changes. As mentioned in the introduction to this section, we can imagine a hypothetical observer attached to the deforming material, who moves and deforms with the material. This observer will see no change in the configuration itself, $\dot{\mathbf{g}}_i = \dot{\mathbf{g}}^i = 0$. However, they will still see changes to vectors and tensors. These changes are measured using the Lie Derivative, which will be seen to be none other than the convected derivative discussed above.

Vectors

First, the **Lie (time) derivative** $\mathrm{L}_v \mathbf{v}$ of a vector $\mathbf{v}$ is the material derivative *holding the deformed basis constant*, that is, Eqns. 2.12.3:

$$\begin{aligned} \mathrm{L}_v^{b}\,\mathbf{v} &= \dot{v}_i \mathbf{g}^i \\ \mathrm{L}_v^{\#}\,\mathbf{v} &= \dot{v}^i \mathbf{g}_i \end{aligned} \tag{2.12.65}$$

Formally, it is defined in terms of the pull-back and push-forward,

$$\boxed{\mathrm{L}_v \mathbf{v} = \chi_*\left(\frac{d}{dt}\left[\chi_*^{-1}(\mathbf{v})\right]\right)} \qquad \textbf{The Lie Time Derivative} \tag{2.12.66}$$

This is illustrated in the Fig. 2.12.3. The spatial vector is first pulled back to the reference configuration, there the differentiation is carried out, where the base vectors are constant, then the vector is pushed forward again to the spatial description.

$$\begin{tikzcd}
\frac{d}{dt}\chi_*^{-1}(\mathbf{v}) \arrow[rr, bend left, "\chi_*"] & & \chi_*\left(\frac{d}{dt}\chi_*^{-1}(\mathbf{v})\right) \\
\chi_*^{-1}(\mathbf{v}) \arrow[u, dashed] & & \mathbf{v} \arrow[ll, bend left, "\chi_*^{-1}"]
\end{tikzcd}$$

Figure 2.12.3: The Lie Derivative

For covariant components, one first pulls back the vector $v_i \mathbf{g}^i$ to $v_i \mathbf{G}^i$, the derivative is taken, $\dot{v}_i \mathbf{G}^i$, and then it is pushed forward to $\dot{v}_i \mathbf{g}^i$, which is consistent with the definition 2.12.65a. The definition 2.12.51 allows one to calculate the Lie derivative in absolute notation: using 2.12.36a, 2.12.33a, 2.11.9,

$$\begin{aligned} \mathrm{L}_v^b \mathbf{v} &= \chi_* \left(\frac{d}{dt} \left[\chi_*^{-1} \left(\mathbf{v} \right)^b \right] \right)^b = \mathbf{F}^{-\mathrm{T}} \left(\frac{d}{dt} \left[\mathbf{F}^{\mathrm{T}} \mathbf{v} \right] \right) \\ &= \mathbf{F}^{-\mathrm{T}} \left(\dot{\mathbf{F}}^{\mathrm{T}} \mathbf{v} + \mathbf{F}^{\mathrm{T}} \dot{\mathbf{v}} \right) \\ &= \mathbf{F}^{-\mathrm{T}} \left(\mathbf{F}^{\mathrm{T}} \mathbf{l}^{\mathrm{T}} \mathbf{v} + \mathbf{F}^{\mathrm{T}} \dot{\mathbf{v}} \right) \\ &= \dot{\mathbf{v}} + \mathbf{l}^{\mathrm{T}} \mathbf{v} \end{aligned} \qquad (2.12.67)$$

The Lie derivative for the contravariant components can be calculated in a similar way, and in summary (these are simply Eqns. 2.12.6): {▲Problem 2}

$$\boxed{\begin{aligned} \mathrm{L}_v^b \mathbf{v} &= \dot{v}_i \mathbf{g}^i = \dot{\mathbf{v}} + \mathbf{l}^{\mathrm{T}} \mathbf{v} \\ \mathrm{L}_v^\# \mathbf{v} &= \dot{v}^i \mathbf{g}_i = \dot{\mathbf{v}} - \mathbf{l} \mathbf{v} \end{aligned}} \qquad \textbf{Lie Derivatives of Vectors} \qquad (2.12.68)$$

Tensors

The material time derivative of a spatial tensor **a** is

$$\begin{aligned} \dot{\mathbf{a}} &= \dot{a}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j + a_{ij} \dot{\mathbf{g}}^i \otimes \mathbf{g}^j + a_{ij} \mathbf{g}^i \otimes \dot{\mathbf{g}}^j \\ &= \dot{a}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j + a^{ij} \dot{\mathbf{g}}_i \otimes \mathbf{g}_j + a^{ij} \mathbf{g}_i \otimes \dot{\mathbf{g}}_j \\ &= \dot{a}^i_{\cdot j} \mathbf{g}_i \otimes \mathbf{g}^j + a^i_{\cdot j} \dot{\mathbf{g}}_i \otimes \mathbf{g}^j + a^i_{\cdot j} \mathbf{g}_i \otimes \dot{\mathbf{g}}^j \\ &= \dot{a}_i^{\cdot j} \mathbf{g}^i \otimes \mathbf{g}_j + a_i^{\cdot j} \dot{\mathbf{g}}^i \otimes \mathbf{g}_j + a_i^{\cdot j} \mathbf{g}^i \otimes \dot{\mathbf{g}}_j \end{aligned} \qquad (2.12.69)$$

The Lie (time) derivative $\mathrm{L}_v \mathbf{a}$ is then

$$\begin{aligned} \mathrm{L}_v^b \mathbf{a} &= \dot{a}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j \\ \mathrm{L}_v^\# \mathbf{a} &= \dot{a}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j \\ \mathrm{L}_v^\backslash \mathbf{a} &= \dot{a}^i_{\cdot j} \mathbf{g}_i \otimes \mathbf{g}^j \\ \mathrm{L}_v^/ \mathbf{a} &= \dot{a}_i^{\cdot j} \mathbf{g}^i \otimes \mathbf{g}_j \end{aligned} \qquad (2.12.70)$$

For example, for covariant components, one first pulls back the tensor $a_{ij} \mathbf{g}^i \otimes \mathbf{g}^j$ to $a_{ij} \mathbf{G}^i \otimes \mathbf{G}^j$, the derivative is taken, $\dot{a}_{ij} \mathbf{G}^i \otimes \mathbf{G}^j$, and then it is pushed forward to $\dot{a}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j$. Thus, using 2.12.42a, 2.12.42a, 2.11.9,

$$
\begin{aligned}
\mathrm{L}_v^b \mathbf{a} &= \chi_* \left(\frac{d}{dt} \left[\chi_*^{-1} (\mathbf{a})^b \right] \right)^b \\
&= \mathbf{F}^{-\mathrm{T}} \left(\dot{\mathbf{F}}^{\mathrm{T}} \mathbf{a} \mathbf{F} + \mathbf{F}^{\mathrm{T}} \dot{\mathbf{a}} \mathbf{F} + \mathbf{F}^{\mathrm{T}} \mathbf{a} \dot{\mathbf{F}} \right) \mathbf{F}^{-1} \\
&= \mathbf{F}^{-\mathrm{T}} \left(\mathbf{F}^{\mathrm{T}} \mathbf{l}^{\mathrm{T}} \mathbf{a} \mathbf{F} + \mathbf{F}^{\mathrm{T}} \dot{\mathbf{a}} \mathbf{F} + \mathbf{F}^{\mathrm{T}} \mathbf{a} \mathbf{l} \mathbf{F} \right) \mathbf{F}^{-1} \\
&= \mathbf{l}^{\mathrm{T}} \mathbf{a} + \dot{\mathbf{a}} + \mathbf{a} \mathbf{l}
\end{aligned}
\tag{2.12.71}
$$

The Lie derivative for the other components can be calculated in a similar way, and in summary (these are Eqns. 2.12.14): {▲Problem 3}

$$
\boxed{
\begin{aligned}
\mathrm{L}_v^b \mathbf{a} &= \dot{a}_{ij} \mathbf{g}^i \otimes \mathbf{g}^j &&= \dot{\mathbf{a}} + \mathbf{l}^{\mathrm{T}} \mathbf{a} + \mathbf{a} \mathbf{l} \\
\mathrm{L}_v^\# \mathbf{a} &= \dot{a}^{ij} \mathbf{g}_i \otimes \mathbf{g}_j &&= \dot{\mathbf{a}} - \mathbf{l} \mathbf{a} - \mathbf{a} \mathbf{l}^{\mathrm{T}} \\
\mathrm{L}_v^\backslash \mathbf{a} &= \dot{a}^i_{\cdot j} \mathbf{g}_i \otimes \mathbf{g}^j &&= \dot{\mathbf{a}} - \mathbf{l} \mathbf{a} + \mathbf{a} \mathbf{l} \\
\mathrm{L}_v^/ \mathbf{a} &= \dot{a}_i^{\cdot j} \mathbf{g}^i \otimes \mathbf{g}_j &&= \dot{\mathbf{a}} + \mathbf{l}^{\mathrm{T}} \mathbf{a} - \mathbf{a} \mathbf{l}^{\mathrm{T}}
\end{aligned}
}
\qquad \textbf{Lie Derivatives of Tensors}
\tag{2.12.72}
$$

The first of these, $\dot{\mathbf{a}} + \mathbf{l}^{\mathrm{T}} \mathbf{a} + \mathbf{a} \mathbf{l}$, is called the **Cotter-Rivlin rate**. The second of these, $\dot{\mathbf{a}} - \mathbf{l} \mathbf{a} - \mathbf{a} \mathbf{l}^{\mathrm{T}}$, is also called the **Oldroyd rate**.

Lie Derivatives of Strain Tensors

From 2.5.18,

$$
\begin{aligned}
\mathbf{d} &= \dot{\mathbf{e}} + \mathbf{l}^{\mathrm{T}} \mathbf{e} + \mathbf{e} \mathbf{l} \\
\dot{\mathbf{b}} - \mathbf{l} \mathbf{b} - \mathbf{b} \mathbf{l}^{\mathrm{T}} &= \mathbf{0}
\end{aligned}
\tag{2.12.73}
$$

and so the Lie derivative of the covariant form of the Euler-Almansi strain is the rate of deformation and the Lie derivative of the contravariant form of the left Cauchy-Green tensor is zero. Further, from 2.12.53a, the Lie derivative of the metric tensor is the push forward of the material time derivative of the right Cauchy-Green strain:

$$
\mathrm{L}_v^b \mathbf{g} = \chi_* \left(\dot{\mathbf{C}} \right)^b, \tag{2.12.74}
$$

Also, directly from 2.11.15,

$$
\mathrm{L}_v^b \mathbf{g} = 2 \mathbf{d} \tag{2.12.75}
$$

Corotational Rates

The Lie derivatives in 2.12.72 were derived using pull-backs and push-forwards between the reference configuration and the current configuration. If, instead, we relate quantities to the rotated intermediate configuration, in other words use **R** instead of **F** in the calculations, we find that, using Eqn. 2.6.1, $\mathbf{\Omega_R} \equiv \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}} = -\mathbf{R}\dot{\mathbf{R}}^{\mathrm{T}}$,

$$\begin{aligned} \mathrm{L}_v\mathbf{a} &= \chi_*\left(\frac{d}{dt}\left[\chi_*^{-1}(\mathbf{a})\right]\right) \\ &= \mathbf{R}\left(\frac{d}{dt}\left[\mathbf{R}^{\mathrm{T}}\mathbf{a}\mathbf{R}\right]\right)\mathbf{R}^{\mathrm{T}} \\ &= \dot{\mathbf{a}} - \mathbf{\Omega_R}\mathbf{a} + \mathbf{a}\mathbf{\Omega_R} \end{aligned} \tag{2.12.76}$$

This is called the **Green-Naghdi rate**.

Rather than pulling back from the intermediate configuration to the reference configuration, we can choose the current configuration to be the reference configuration. Rotating from this configuration (see section 2.6.3), $\mathbf{\Omega_R} = \mathbf{w}$, the spin tensor, and one obtains the **Jaumann rate**, $\dot{\mathbf{a}} - \mathbf{wa} + \mathbf{aw}$.

Lie Derivatives and Objective Rates

The concept of objectivity was discussed in section 2.8. Essentially, if two observers are rotating relative to each other with rotation $\mathbf{Q}(t)$ and both are observing some spatial tensor, **T** as measured by one observer and $\mathbf{T}^*$ as measured by the other, then this tensor is objective provided $\mathbf{T}^* = \mathbf{QTQ}^{\mathrm{T}}$ for all **Q**, i.e. the measurement of the deformation would be independent of the observer. One of the most important uses of the Lie derivative is that *Lie derivatives of objective spatial tensors are objective spatial tensors*. Thus the rates given in 2.12.72 are all objective.

For example, suppose we have an objective spatial tensor **a**, i.e. so that $\mathbf{a}^* = \mathbf{QaQ}^{\mathrm{T}}$. The velocity gradient is not objective, and instead satisfies the relation 2.8.27: $\mathbf{l}^* = \mathbf{QlQ}^{\mathrm{T}} + \dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}$. Using the properties of the transpose, the orthogonality of **Q**, and the identity $\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\dot{\mathbf{Q}}^{\mathrm{T}}$, one has for Eqns. 2.12.72a,b,

$$
\begin{aligned}
\left(\mathrm{L}_v^b \mathbf{a}\right)^* &= \dot{\overline{\mathbf{a}^*}} + \mathbf{l}^{*\mathrm{T}}\mathbf{a}^* + \mathbf{a}^*\mathbf{l}^* \\
&= \dot{\overline{\mathbf{QaQ}^\mathrm{T}}} + \left(\mathbf{QlQ}^\mathrm{T} + \dot{\mathbf{Q}}\mathbf{Q}^\mathrm{T}\right)^\mathrm{T}\left(\mathbf{QaQ}^\mathrm{T}\right) + \left(\mathbf{QaQ}^\mathrm{T}\right)\left(\mathbf{QlQ}^\mathrm{T} + \dot{\mathbf{Q}}\mathbf{Q}^\mathrm{T}\right) \\
&= \dot{\mathbf{Q}}\mathbf{aQ}^\mathrm{T} + \mathbf{Q}\dot{\mathbf{a}}\mathbf{Q}^\mathrm{T} + \mathbf{Qa}\dot{\mathbf{Q}}^\mathrm{T} + \mathbf{Ql}^\mathrm{T}\mathbf{Q}^\mathrm{T}\mathbf{QaQ}^\mathrm{T} + \mathbf{Q}\dot{\mathbf{Q}}^\mathrm{T}\mathbf{QaQ}^\mathrm{T} \\
&\qquad + \mathbf{QaQ}^\mathrm{T}\mathbf{QlQ}^\mathrm{T} + \mathbf{QaQ}^\mathrm{T}\dot{\mathbf{Q}}\mathbf{Q}^\mathrm{T} \\
&= \mathbf{Q}\left(\dot{\mathbf{a}} + \mathbf{l}^\mathrm{T}\mathbf{a} + \mathbf{al}\right)\mathbf{Q}^\mathrm{T} \\
\left(\mathrm{L}_v^\# \mathbf{a}\right)^* &= \dot{\overline{\mathbf{a}^*}} - \mathbf{l}^*\mathbf{a}^* - \mathbf{a}^*\mathbf{l}^{*\mathrm{T}} \\
&= \dot{\overline{\mathbf{QaQ}^\mathrm{T}}} - \left(\mathbf{QlQ}^\mathrm{T} + \dot{\mathbf{Q}}\mathbf{Q}^\mathrm{T}\right)\left(\mathbf{QaQ}^\mathrm{T}\right) - \left(\mathbf{QaQ}^\mathrm{T}\right)\left(\mathbf{QlQ}^\mathrm{T} + \dot{\mathbf{Q}}\mathbf{Q}^\mathrm{T}\right)^\mathrm{T} \\
&= \dot{\mathbf{Q}}\mathbf{aQ}^\mathrm{T} + \mathbf{Q}\dot{\mathbf{a}}\mathbf{Q}^\mathrm{T} + \mathbf{Qa}\dot{\mathbf{Q}}^\mathrm{T} - \mathbf{QlQ}^\mathrm{T}\mathbf{QaQ}^\mathrm{T} - \dot{\mathbf{Q}}\mathbf{Q}^\mathrm{T}\mathbf{QaQ}^\mathrm{T} \\
&\qquad - \mathbf{QaQ}^\mathrm{T}\mathbf{Ql}^\mathrm{T}\mathbf{Q}^\mathrm{T} - \mathbf{QaQ}^\mathrm{T}\mathbf{Q}\dot{\mathbf{Q}}^\mathrm{T} \\
&= \mathbf{Q}\left(\dot{\mathbf{a}} - \mathbf{la} - \mathbf{al}^\mathrm{T}\right)\mathbf{Q}^\mathrm{T}
\end{aligned}
\tag{2.12.77}
$$

showing that these rates are indeed objective.

Further, any linear combination of them is objective, for example,

$$
\frac{1}{2}\left[\left(\dot{\mathbf{a}} + \mathbf{l}^\mathrm{T}\mathbf{a} + \mathbf{al}\right) + \left(\dot{\mathbf{a}} - \mathbf{la} - \mathbf{al}^\mathrm{T}\right)\right] = \dot{\mathbf{a}} + \frac{1}{2}\left[-\left(\mathbf{l} - \mathbf{l}^\mathrm{T}\right)\mathbf{a} + \mathbf{a}\left(\mathbf{l} - \mathbf{l}^\mathrm{T}\right)\right] = \dot{\mathbf{a}} - \mathbf{wa} + \mathbf{aw} \tag{2.12.78}
$$

is objective, provided **a** is. This is the **Jaumann rate** introduced in Eqn. 2.8.36 and mentioned after Eqn. 2.12.76 above. Further, as mentioned after Eqn. 2.12.72, the **Cotter-Rivlin rate** of Eqn. 2.8.37 is equivalent to $\mathrm{L}_v^b \mathbf{a}$.

The Lie Derivative and the Directional Derivative

Recall that the material time derivative of a tensor can be written in terms of the directional derivative, §2.6.5. Hence the Lie derivative can also be expressed as

$$
\mathrm{L}_v \mathbf{T} = \chi_*\left(\partial_\mathbf{f}\left(\chi_*^{-1}(\mathbf{T})\right)[\mathbf{v}]\right) \tag{2.12.79}
$$

and hence the subscript v on the L. Thus one can say that the Lie derivative is the push forward of the directional derivative of the material field $\chi_*^{-1}(\mathbf{T})$ in the direction of the velocity vector.

2.12.4 Problems

1. Eqns. 2.12.30 follow immediately from 2.12.29. However, use Eqns. 2.12.40, 2.12.42, i.e. $\chi_*(\mathbf{A})^b = \mathbf{F}^{-\mathrm{T}}\mathbf{A}\mathbf{F}^{-1}$, etc., directly, to verify relations 2.12.45.
2. Derive the Lie derivatives of a vector **v**, Eqns. 2.12.68.
3. Derive the Lie derivatives of a tensor **a**, Eqns. 2.12.72.

2.13 Variation and Linearisation of Kinematic Tensors

2.13.1 The Variation of Kinematic Tensors

The Variation

In this section is reviewed the concept of the variation, introduced in Part I, §8.5.

The **variation** is defined as follows: consider a function $\mathbf{u}(\mathbf{x})$, with $\mathbf{u}^*(\mathbf{x})$ a second function which is at most infinitesimally different from $\mathbf{u}(\mathbf{x})$ at every point ***x***, Fig. 2.13.1

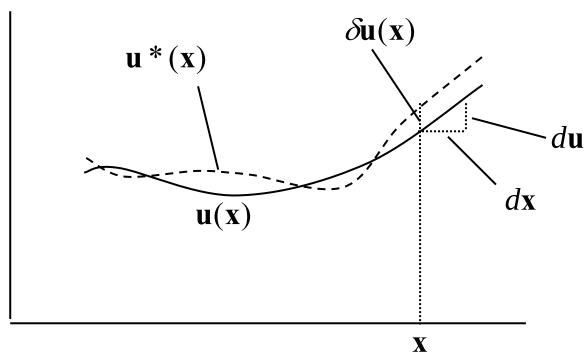


Figure 2.13.1: the variation

Then define

$$\boxed{\delta\mathbf{u} = \mathbf{u}^*(\mathbf{x}) - \mathbf{u}(\mathbf{x})} \qquad \textbf{The Variation} \qquad (2.13.1)$$

The operator δ is called the **variation symbol** and $\delta\mathbf{u}$ is called the variation of $\mathbf{u}(\mathbf{x})$.

The variation of $\mathbf{u}(\mathbf{x})$ is understood to represent an infinitesimal change in the function *at **x***. Note from the figure that a variation $\delta\mathbf{u}$ of a function **u** is different to a differential $d\mathbf{u}$. The ordinary differentiation gives a measure of the change of a function resulting from a specified change in the *independent* variable (in this case **x**). Also, note that the independent variable does not participate in the variation process; the variation operator imparts an infinitesimal change to the function **u** at some *fixed* **x** – formally, one can write this as $\delta\mathbf{x} = 0$.

The Commutative Properties of the variation operator

(1) $$\frac{d}{d\mathbf{x}}\delta\mathbf{u} = \delta\frac{d\mathbf{u}}{d\mathbf{x}} \qquad (2.13.2)$$

Proof:

$$\delta \frac{d\mathbf{u}}{d\mathbf{x}} = \left(\frac{d\mathbf{u}}{d\mathbf{x}}\right)^* - \frac{d\mathbf{u}}{d\mathbf{x}} = \frac{d\mathbf{u}^*}{d\mathbf{x}} - \frac{d\mathbf{u}}{d\mathbf{x}} = \frac{d(\mathbf{u}^* - \mathbf{u})}{d\mathbf{x}} = \frac{d}{d\mathbf{x}}\left(\delta \mathbf{u}(\mathbf{x})\right)$$

(2) $$\delta \int_{x_1}^{x_2} \mathbf{u}(\mathbf{x}) d\mathbf{x} = \int_{x_1}^{x_2} \delta \mathbf{u}(\mathbf{x}) d\mathbf{x} \tag{2.13.3}$$

Proof:

$$\delta \int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{u}(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{u}^*(\mathbf{x}) d\mathbf{x} - \int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{u}(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{x}_1}^{\mathbf{x}_2} [\mathbf{u}^*(\mathbf{x}) - \mathbf{u}(\mathbf{x})] d\mathbf{x} = \int_{\mathbf{x}_1}^{\mathbf{x}_2} \delta \mathbf{u}(\mathbf{x}) d\mathbf{x}$$

Variation of a Function

Consider **A**, a scalar-, vector-, or tensor-valued function of $\mathbf{u}$, $\mathbf{A}(\mathbf{u})$. When we apply a variation to **u**, $\delta\mathbf{u}$, **A** changes to $\mathbf{A}(\mathbf{u} + \delta\mathbf{u})$. The variation of **A** is then defined as

$$\delta \mathbf{A}(\mathbf{u}, \delta \mathbf{u}) = \mathbf{A}(\mathbf{u} + \delta \mathbf{u}) - \mathbf{A}(\mathbf{u}) \tag{2.13.4}$$

(in the limit as $\delta\mathbf{u} \to 0$). This can be expressed using the concept of the directional derivative in the usual way (see §1.6.11): consider the function $\mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u})$, so that $\mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u})_{\varepsilon=0} = \mathbf{A}(\mathbf{u})$ and $\mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u})_{\varepsilon=1} = \mathbf{A}(\mathbf{u} + \delta\mathbf{u})$. A Taylor expansion gives $\mathbf{A}(\varepsilon) = \mathbf{A}(0) + \varepsilon(d\mathbf{A}/d\varepsilon)_{\varepsilon=0} + \cdots$, or

$$\mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u}) = \mathbf{A}(\mathbf{u}) + \varepsilon\left(\frac{d}{d\varepsilon}\mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u})\right)_{\varepsilon=0} + \cdots \tag{2.13.5}$$

Setting $\varepsilon = 1$ then gives Eqn. 2.13.4; thus

$$\mathbf{A}(\mathbf{u} + \delta\mathbf{u}) \approx \mathbf{A}(\mathbf{u}) + \partial_{\mathbf{u}}\mathbf{A}[\delta\mathbf{u}] \tag{2.13.6}$$

where $\partial_{\mathbf{u}}\mathbf{A}[\delta\mathbf{u}]$ is the directional derivative of **A** in the direction $\delta\mathbf{u}$; the directional derivative in this context is the variation of **A**:

$$\delta\mathbf{A}(\mathbf{u}, \delta\mathbf{u}) \equiv \partial_{\mathbf{u}}\mathbf{A}[\delta\mathbf{u}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u}) \tag{2.13.7}$$

For example, consider the scalar function $\phi = \mathbf{P} : \mathbf{E}$, where $\mathbf{P}$ and $\mathbf{E}$ are second order tensors. Then

$$\delta\phi(\mathbf{E}, \delta\mathbf{E}) \equiv \partial_{\mathbf{E}}\phi[\delta\mathbf{E}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{P} : (\mathbf{E} + \varepsilon\delta\mathbf{E}) = \mathbf{P} : \delta\mathbf{E} \tag{2.13.8}$$

The **second variation** is defined as

$$\delta^2\mathbf{A} = \delta(\delta\mathbf{A}) = \partial_{\mathbf{u}}\delta\mathbf{A}[\delta\mathbf{u}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \delta\mathbf{A}(\mathbf{u} + \varepsilon\delta\mathbf{u}) \tag{2.13.9}$$

For example, for a scalar function $\phi(\mathbf{u})$ of a vector $\mathbf{u}$, the chain rule and Eqn. 2.13.2 give

$$\begin{aligned} \delta\varphi(\mathbf{u}, \delta\mathbf{u}) &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \varphi(\mathbf{u} + \varepsilon\delta\mathbf{u}) = \left.\frac{d\varphi(\mathbf{u} + \varepsilon\delta\mathbf{u})}{d(\mathbf{u} + \varepsilon\delta\mathbf{u})}\right|_{\varepsilon=0} \frac{d(\mathbf{u} + \varepsilon\delta\mathbf{u})}{d\varepsilon} = \frac{\partial\varphi}{\partial\mathbf{u}} \cdot \delta\mathbf{u} \\ \delta^2\varphi &= \frac{\partial\delta\varphi}{\partial\mathbf{u}} \cdot \delta\mathbf{u} = \left(\delta\frac{\partial\varphi}{\partial\mathbf{u}}\right) \cdot \delta\mathbf{u} = \left(\frac{\partial^2\varphi}{\partial\mathbf{u}\partial\mathbf{u}}\delta\mathbf{u}\right) \cdot \delta\mathbf{u} = \delta\mathbf{u}\frac{\partial^2\varphi}{\partial\mathbf{u}\partial\mathbf{u}}\delta\mathbf{u} \end{aligned} \tag{2.13.10}$$

Variation of Functions of the Displacement

In what follows is discussed the change (variation) in functions $\mathbf{A}(\mathbf{u})$ when the displacement (or velocity) fields undergo a variation. These ideas are useful in formulating variational principles of mechanics (see, for example, §3.9).

Shown in Fig. 2.13.2 is the current configuration frozen at some instant in time. The displacement field is then allowed to undergo a variation $\delta\mathbf{u}$. This change to the displacement field evidently changes kinematic tensors, and these changes are now investigated. Note that this variation to the displacement induces a variation to $\mathbf{x}$, $\delta\mathbf{x}$, but $\mathbf{X}$ remains unchanged, $\delta\mathbf{X} = 0$.

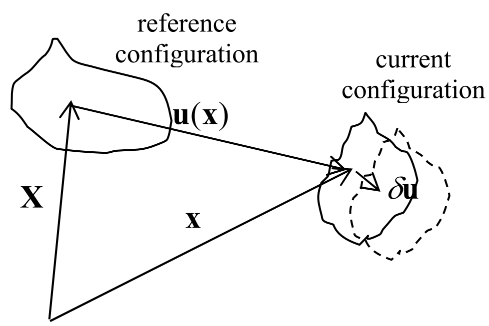


Figure 2.13.2: a variation of the displacement

To evaluate the variation of the deformation gradient $\mathbf{F}$, $\delta\mathbf{F}(\mathbf{u},\delta\mathbf{u})$, where $\mathbf{u}$ is the displacement field, note that $\mathbf{u}=\mathbf{x}-\mathbf{X}$ and Eqn. 2.2.43, $\mathbf{F}(\mathbf{u})=\text{Grad}\mathbf{u}+\mathbf{I}$. One has, from the definition 2.13.7,

$$\begin{aligned}\delta\mathbf{F}(\mathbf{u},\delta\mathbf{u})=\partial_{\mathbf{u}}\mathbf{F}[\delta\mathbf{u}]&=\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\mathbf{F}(\mathbf{u}+\varepsilon\delta\mathbf{u})\\&=\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\left[\mathbf{F}(\mathbf{u})+\varepsilon\text{Grad}(\delta\mathbf{u})+\mathbf{I}\right]\\&=\text{Grad}(\delta\mathbf{u})\end{aligned}\tag{2.13.11}$$

Noting the first commutative property of the variation, 2.13.2, this can also be expressed as

$$\delta\mathbf{F}(\mathbf{u},\delta\mathbf{u})=\delta(\text{Grad}\mathbf{u})\tag{2.13.12}$$

Note that $\delta\mathbf{u}$ is completely independent of the function $\mathbf{u}$.

Here are some other examples, involving the inverse deformation gradient, the Green-Lagrange strain, the inverse right Cauchy-Green strain and the spatial line element: {▲Problem 1-3}

$$\begin{aligned}\delta\mathbf{F}^{-1}&=-\mathbf{F}^{-1}\text{grad}\delta\mathbf{u}\\\delta\mathbf{E}&=\mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F}\\\delta\mathbf{C}^{-1}&=-2\mathbf{F}^{-1}\boldsymbol{\varepsilon}\mathbf{F}^{-\mathrm{T}}\end{aligned}\tag{2.13.13}$$

where $\boldsymbol{\varepsilon}$ is the small strain tensor, Eqn. 2.2.48.

One also has, using the chain rule for the directional derivative, Eqn. 1.15.28, the directional derivative for the determinant, Eqn. 1.15.32, the trace relation 1.10.10e, Eqn. 2.2.8b,

$$\begin{aligned}\delta J(\mathbf{u},\delta\mathbf{u})&=\delta\det\mathbf{F}(\mathbf{u},\delta\mathbf{u})\\&=\partial_{\mathbf{u}}\det\mathbf{F}[\delta\mathbf{u}]\\&=\partial_{\mathbf{F}}\det\mathbf{F}[\partial_{\mathbf{u}}\mathbf{F}[\delta\mathbf{u}]]\\&=\partial_{\mathbf{F}}\det\mathbf{F}[\text{Grad}(\delta\mathbf{u})]\\&=\det\mathbf{F}\left[\mathbf{F}^{-\mathrm{T}}:\text{Grad}(\delta\mathbf{u})\right]\\&=J\text{tr}\left(\text{Grad}(\delta\mathbf{u})\mathbf{F}^{-1}\right)\\&=J\text{tr}(\text{grad}(\delta\mathbf{u}))\\&=J\,\text{div}(\delta\mathbf{u})\end{aligned}\tag{2.13.14}$$

Example

To put some of the above concepts into a simple and less abstract setting, consider the following scenario: a bar over $0 \le \mathbf{X} \le 1$ is extended, as illustrated in Fig. 2.13.3, according to:

$$\begin{aligned} \mathbf{x} &= 2\mathbf{X}^2 + 3 \\ \mathbf{X} &= \sqrt{\tfrac{1}{2}(\mathbf{x}-3)} \end{aligned} \tag{2.13.15}$$

The deformation gradient is

$$\mathbf{F} = \mathrm{Grad}\mathbf{x} = 4\mathbf{X} \tag{2.13.16}$$

So, for example, in the initial configuration (*A*), an infinitesimal line element at $\mathbf{X} = 0$ does not stretch ($\mathbf{F} = 0$) whereas a line element at $\mathbf{X} = 1$ stretches by 4.

The inverse deformation gradient is

$$\mathbf{F}^{-1} = \mathrm{grad}\mathbf{X} = \frac{1}{\sqrt{8(\mathbf{x}-3)}} \tag{2.13.17}$$

This implies that, in the current configuration (*B*), an infinitesimal line element at $\mathbf{x} = 3$ is the same size as its counterpart in the initial configuration ($\mathbf{F}^{-1} = 0$) whereas a line element at $\mathbf{x} = 5$ shrinks by a factor of 4 when returning to the initial configuration

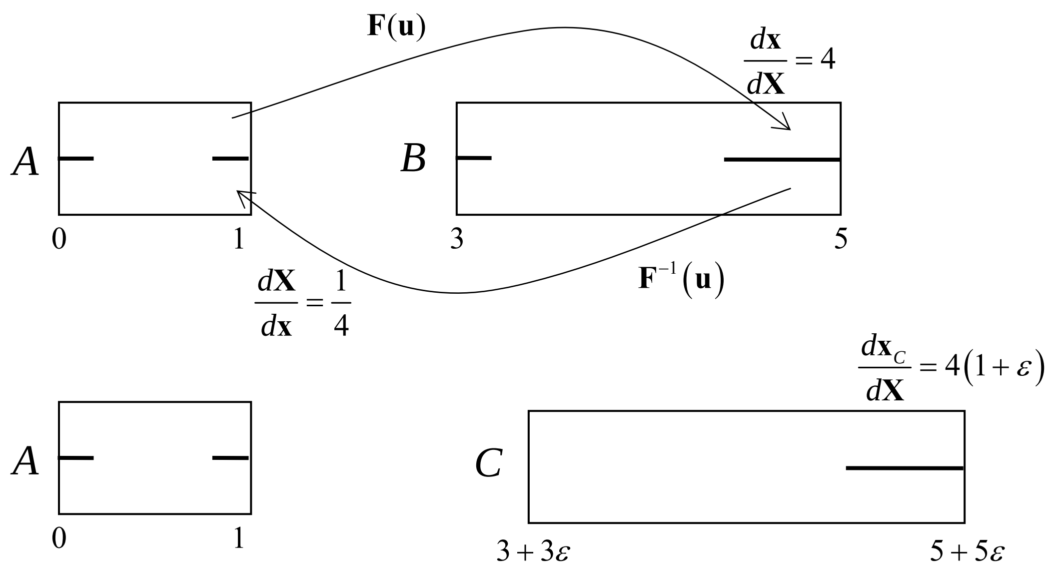


Figure 2.13.3: a motion and a variation

Now introduce a variation, which moves the bar from configuration *B* to configuration *C*:

$$\delta \mathbf{u} = \varepsilon \mathbf{x} = \varepsilon\left(2\mathbf{X}^2 + 3\right) \tag{2.13.18}$$

The point at 3 moves to $3+3\varepsilon$ and the point at 5 moves to $5+5\varepsilon$. (This variation happens to be a simple linear function of **x**, but it can be anything for our purposes here.)

The variation is plotted below as a function of **X** and **x**.

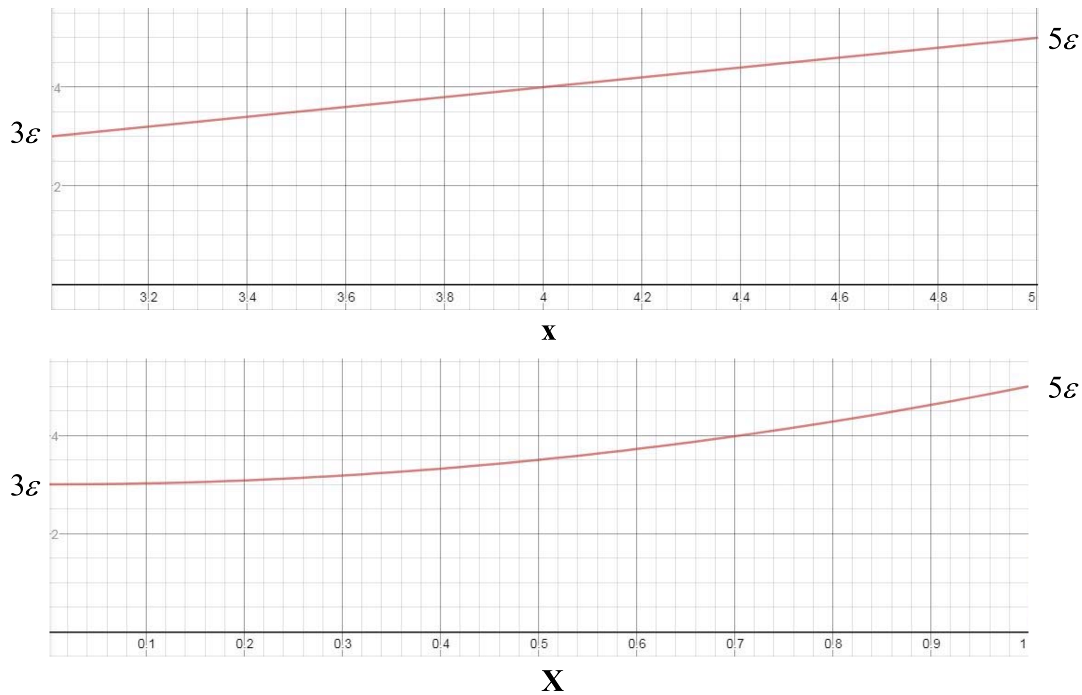


Figure 2.13.4: the variation as a function of x and X

Differentiating Eqns. 2.13.19, the gradients of the variations are

$$\begin{aligned}\operatorname{Grad}(\delta \mathbf{u}) &= \varepsilon(4\mathbf{X}) \\ \operatorname{grad}(\delta \mathbf{u}) &= \varepsilon\end{aligned} \tag{2.13.20}$$

which are the slopes in Figure 2.13.4.

To calculate the **F** associated with the new variation configuration, i.e. $\mathbf{F}(\mathbf{u}+\delta\mathbf{u})$, note that points **X** have now moved to:

$$2\mathbf{X}^2 + 3 + \varepsilon\left(2\mathbf{X}^2 + 3\right) \tag{2.13.21}$$

and so

$$\mathbf{F}(\mathbf{u}+\delta\mathbf{u}) = \mathrm{Grad}\big((1+\varepsilon)(2\mathbf{X}^2+3)\big) = 4\mathbf{X}+\varepsilon 4\mathbf{X} \tag{2.13.22}$$

This says that an infinitesimal line element at $\mathbf{X}=0$ does not stretch when moving to configuration *C* ($\mathbf{F}=0$) whereas a line element at $\mathbf{X}=1$ stretches by $4+4\varepsilon$.

Subtracting Eqn. 2.13.17 form Eqn. 2.13.22:

$$\delta\mathbf{F} = \mathbf{F}(\mathbf{u}+\delta\mathbf{u}) - \mathbf{F}(\mathbf{u}) = \varepsilon(4\mathbf{X}) \tag{2.13.23}$$

From Eqn. 2.13.20, this verifies Eqn. 2.13.11, that

$$\delta\mathbf{F} = \mathrm{Grad}(\delta\mathbf{u}) \tag{2.13.24}$$

We could also calculate the variation of **F** by moving directly from configuration *B* to configuration *C*. The movement of the particles from *B* to *C* is given by Eqn. 2.13.19: $\varepsilon(2\mathbf{X}^2+3)$ and so, based on this motion, $\delta\mathbf{F} = \mathrm{Grad}\big(\varepsilon(2\mathbf{X}^2+3)\big) = \varepsilon(4\mathbf{X})$.

To calculate the $\mathbf{F}^{-1}$ associated with the new variation configuration, i.e. $\mathbf{F}^{-1}(\mathbf{u}+\delta\mathbf{u})$, note that the "new" current position **x** is (Eqn. 2.13.21):

$$\begin{aligned} \mathbf{x}_C &= 2\mathbf{X}^2+3+\varepsilon(2\mathbf{X}^2+3) \\ \to \mathbf{X} &= \sqrt{\frac{1}{2}\left(\frac{\mathbf{x}_C}{1+\varepsilon}-3\right)} \end{aligned} \tag{2.13.25}$$

This means that the point $3+3\varepsilon$ in configuration *C* corresponds to $\mathbf{X}=0$ and the point $5+5\varepsilon$ corresponds to $\mathbf{X}=1$. Then,

$$\mathbf{F}^{-1}(\mathbf{u}+\delta\mathbf{u}) = \frac{d}{d\mathbf{x}_C}\sqrt{\frac{1}{2}\left(\frac{\mathbf{x}_C}{1+\varepsilon}-3\right)} = \frac{1}{1+\varepsilon}\frac{1}{\sqrt{8\left(\frac{\mathbf{x}_C}{1+\varepsilon}-3\right)}} \tag{2.13.26}$$

So an element at the point $3+3\varepsilon$ in configuration *C* does not change in size as it is mapped back to the initial configuration, whereas an element at the point $5+5\varepsilon$ shrinks back to the initial configuration by a factor of $1/(4+4\varepsilon)$, as indicated in Fig. 2.13.3.

Alternatively, since $\mathbf{x}_C = \mathbf{x}+\varepsilon\mathbf{x}$, this can be written as

$$\mathbf{F}^{-1}(\mathbf{u}+\delta\mathbf{u}) = \frac{1}{1+\varepsilon}\frac{1}{\sqrt{8(\mathbf{x}-3)}} \tag{2.13.27}$$

Subtracting Eqn. 2.13.18 from Eqn. 2.13.27, the variation of the inverse deformation gradient is then

$$\begin{aligned}\delta\mathbf{F}^{-1}(\mathbf{u}) = \mathbf{F}^{-1}(\mathbf{u}+\delta\mathbf{u}) - \mathbf{F}^{-1}(\mathbf{u}) &= \frac{1}{1+\varepsilon}\frac{1}{\sqrt{8(\mathbf{x}-3)}} - \frac{1}{\sqrt{8(\mathbf{x}-3)}} \\ &= -\frac{\varepsilon}{1+\varepsilon}\frac{1}{\sqrt{8(\mathbf{x}-3)}}\end{aligned} \tag{2.13.28}$$

Using a series expansion, $(1+\varepsilon)^{-1} = 1-\varepsilon+\varepsilon^2-\ldots$, for small ε (neglecting terms of order ε^2),

$$\delta\mathbf{F}^{-1}(\mathbf{u}) = -\varepsilon\frac{1}{\sqrt{8(\mathbf{x}-3)}} \tag{2.13.29}$$

From Eqns. 2.13.18 and 2.13.20, this verifies the relation 2.13.13:

$$\delta\mathbf{F}^{-1}(\mathbf{u}) = -\mathbf{F}^{-1}\mathrm{grad}(\delta u) \tag{2.13.30}$$

A formula for the inverse deformation gradient is $\mathbf{F}^{-1} = \mathbf{I} - \mathrm{grad}\mathbf{u}$. However, note that $\mathbf{F}^{-1}(\mathbf{u}+\delta\mathbf{u}) \neq \mathbf{I} - \partial\mathbf{u}/\partial\mathbf{x}$, but that $\mathbf{F}^{-1}(\mathbf{u}+\delta\mathbf{u}) = \mathbf{I} - \partial\mathbf{u}/\partial\mathbf{x}_C$.

The Lie Variation

The **Lie-variation** is defined for *spatial* vectors and tensors as a variation holding the deformed basis constant. For example,

$$\delta_{\mathrm{L}}^{\mathrm{b}}\mathbf{a} = \delta a_{ij}\mathbf{g}^i \otimes \mathbf{g}^j \tag{2.13.31}$$

The object is first pulled-back, the variation is then taken and finally a push-forward is carried out. For example, analogous to 2.12.66,

$$\delta_{\mathrm{L}}\mathbf{a}(\mathbf{u},\delta\mathbf{u}) \equiv \chi_*\left(\partial_{\mathbf{u}}\left(\chi_*^{-1}(\mathbf{a})\right)[\delta\mathbf{u}]\right) \tag{2.13.32}$$

For example, consider the Lie-variation of the Euler-Almansi strain $\mathbf{e}$. First, from 2.12.56b, $\chi_{-1}^*(\mathbf{e})^b = \mathbf{E}$. Then 2.13.13b gives $\partial_{\mathbf{u}}\left(\chi_{-1}^*(\mathbf{e})^b\right)[\delta\mathbf{u}] = \delta\mathbf{E} = \mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F}$. From 2.12.40a,

$$\delta_{\mathrm{L}}\mathbf{e}(\mathbf{u},\delta\mathbf{u}) = \chi_*\left(\partial_{\mathbf{u}}\left(\chi_{-1}^*(\mathbf{e})^b\right)[\delta\mathbf{u}]\right)^b = \chi_*\left(\mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F}\right)^b = \delta\boldsymbol{\varepsilon} \tag{2.13.33}$$

2.13.2 Linearisation of Kinematic Functions

Linearisation of a Function

As for the variation, consider **A**, a scalar-, vector-, or tensor-valued function of $\mathbf{u}$. If **u** undergoes an increment $\Delta\mathbf{u}$, then, analogous to 2.13.4,

$$\mathbf{A}(\mathbf{u}+\Delta\mathbf{u}) \approx \mathbf{A}(\mathbf{u}) + \partial_{\mathbf{u}}\mathbf{A}[\Delta\mathbf{u}] \tag{1.13.34}$$

The directional derivative $\partial_{\mathbf{u}}\mathbf{A}[\Delta\mathbf{u}]$ in this context is also denoted by $\Delta\mathbf{A}(\mathbf{u},\Delta\mathbf{u})$. The **linearization** of **A** with respect to **u** is defined to be

$$\mathrm{L}\ \mathbf{A}(\mathbf{u},\Delta\mathbf{u}) = \mathbf{A}(\mathbf{u}) + \Delta\mathbf{A}(\mathbf{u},\Delta\mathbf{u}) \tag{1.13.35}$$

Using exactly the same method of calculation as was used for the variations above, the linearization of **F** and **E**, for example, are

$$\begin{aligned} \mathrm{L}\,\mathbf{F}(\mathbf{u},\Delta\mathbf{u}) &= \mathbf{F}(\mathbf{u}) + \partial_{\mathbf{u}}\mathbf{F}[\Delta\mathbf{u}] = \mathbf{F} + \mathrm{Grad}\Delta\mathbf{u} \\ \mathrm{L}\,\mathbf{E}(\mathbf{u},\Delta\mathbf{u}) &= \mathbf{E}(\mathbf{u}) + \partial_{\mathbf{u}}\mathbf{E}[\Delta\mathbf{u}] = \mathbf{E} + \mathbf{F}^{\mathrm{T}}\Delta\boldsymbol{\varepsilon}\mathbf{F} \end{aligned} \tag{2.13.36}$$

where $\Delta\boldsymbol{\varepsilon} = \frac{1}{2}\left((\mathrm{grad}\Delta\mathbf{u})^{\mathrm{T}} + (\mathrm{grad}\Delta\mathbf{u})\right)$ is the linearised small strain tensor $\boldsymbol{\varepsilon}$.

Linearisation of Variations of a Function

One can also linearise the variation of a function. For example,

$$\mathrm{L}\ \delta\mathbf{A}(\mathbf{u},\Delta\mathbf{u}) = \delta\mathbf{A}(\mathbf{u},\delta\mathbf{u}) + \Delta\delta\mathbf{A}(\mathbf{u},\Delta\mathbf{u}) \tag{2.13.37}$$

The second term here is the directional derivative

$$\begin{aligned} \Delta\delta\mathbf{A}[\mathbf{u},\Delta\mathbf{u}] &= \partial_{\mathbf{u}}\delta\mathbf{A}[\Delta\mathbf{u}] \\ &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \delta\mathbf{A}(\mathbf{u}+\varepsilon\Delta\mathbf{u}) \end{aligned} \tag{2.13.38}$$

This leads to an expression similar to $\delta^2\mathbf{A}$. For example, for a scalar function $\phi(\mathbf{u})$ of a vector **u**,

$$\Delta\delta\phi = \frac{\partial\delta\phi}{\partial\mathbf{u}}\cdot\Delta\mathbf{u} = \Delta\mathbf{u}\frac{\partial^2\phi}{\partial\mathbf{u}\partial\mathbf{u}}\delta\mathbf{u} \tag{2.13.39}$$

Consider now the virtual Green-Lagrange strain, 2.13.11b, $\delta\mathbf{E} = \mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F}$. To carry out the linearization of $\delta\mathbf{E}$, it is convenient to first write it in the form

$$\begin{aligned}\delta\mathbf{E} &= \mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F} \\ &= \tfrac{1}{2}\mathbf{F}^{\mathrm{T}}\left[(\mathrm{grad}\,\delta\mathbf{u})^{\mathrm{T}} + \mathrm{grad}\,\delta\mathbf{u}\right]\mathbf{F} \\ &= \tfrac{1}{2}\left[(\mathrm{Grad}\,\delta\mathbf{u})^{\mathrm{T}}\mathbf{F} + \mathbf{F}^{\mathrm{T}}\mathrm{Grad}\,\delta\mathbf{u}\right]\end{aligned} \tag{2.13.40}$$

Then

$$\Delta\delta\mathbf{E} = \partial_{\mathbf{u}}\delta\mathbf{E}[\Delta\mathbf{u}] = \partial_{\mathbf{u}}\left\{\tfrac{1}{2}\left[(\mathrm{Grad}\,\delta\mathbf{u})^{\mathrm{T}}\mathbf{F} + \mathbf{F}^{\mathrm{T}}\mathrm{Grad}\,\delta\mathbf{u}\right]\right\}[\Delta\mathbf{u}] \tag{2.13.41}$$

Recall that the variation $\delta\mathbf{u}$ is *independent* of $\mathbf{u}$; this equation is being linearised with respect to $\mathbf{u}$, and $\delta\mathbf{u}$ is unaffected by the linearization (see Fig. 2.13.3 below). However, the motion, and in particular $\mathbf{F}$, *are* affected by the increment in $\mathbf{u}$. Thus {▲Problem 4}

$$\Delta\delta\mathbf{E} = \mathrm{sym}\left((\mathrm{Grad}\,\Delta\mathbf{u})^{\mathrm{T}}\mathrm{Grad}\,\delta\mathbf{u}\right) \tag{2.13.42}$$

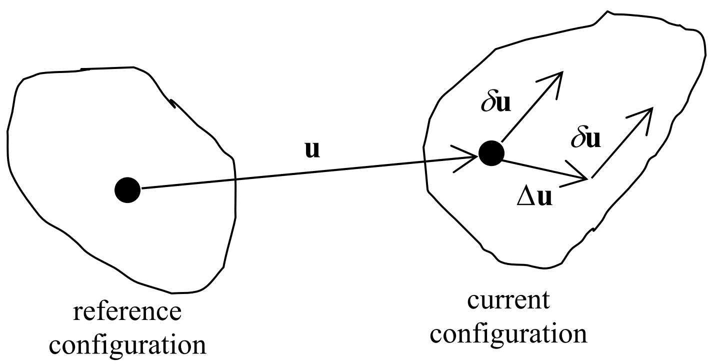


Figure 2.13.3: linearisation

As with the variational operator, one can define the linearization of a spatial tensor as involving a pull back, followed by the directional derivative, and finally the push forward operation. Thus

$$\Delta\mathbf{a}(\mathbf{u},\Delta\mathbf{u}) \equiv \chi_*\left(\partial_{\mathbf{u}}\left(\chi_*^{-1}(\mathbf{a})\right)[\Delta\mathbf{u}]\right) \tag{2.13.43}$$

2.13.3 Problems

1. Use Eqn. 2.2.22, $\mathbf{E} = \frac{1}{2}\left(\mathbf{F}^{\mathrm{T}}\mathbf{F} - \mathbf{I}\right)$, Eqn. 2.13.11, $\delta\mathbf{F}(\mathbf{u}, \delta\mathbf{u}) = \mathrm{Grad}(\delta\mathbf{u})$, and Eqn. 2.2.8b, $\mathrm{grad}\mathbf{v} = (\mathrm{Grad}\mathbf{v})\mathbf{F}^{-1}$, to show that $\delta\mathbf{E} = \mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F}$, where $\boldsymbol{\varepsilon}$ is the small strain tensor, Eqn. 2.2.48.
2. Use 2.13.11 to show that the variation of the inverse deformation gradient $\mathbf{F}^{-1}$ is $\delta\mathbf{F}^{-1} = -\mathbf{F}^{-1}\mathrm{grad}\delta\mathbf{u}$. [Hint: differente the relation $\mathbf{F}^{-1}\mathbf{F} = \mathbf{I}$ by the product rule and then use the relation $\mathrm{grad}\mathbf{v} = (\mathrm{Grad}\mathbf{v})\mathbf{F}^{-1}$ for vector **v**.]
3. Use the definition $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$ to show that $\delta\mathbf{C}^{-1} = -2\mathbf{F}^{-1}\boldsymbol{\varepsilon}\mathbf{F}^{-\mathrm{T}}$.
4. Use the relation $\mathrm{sym}\mathbf{A} = \frac{1}{2}\left(\mathbf{A}^{\mathrm{T}} + \mathbf{A}\right)$ to show that
$$\Delta\delta\mathbf{E} = \partial_{\mathbf{u}}\left\{\tfrac{1}{2}\left[(\mathrm{Grad}\delta\mathbf{u})^{\mathrm{T}}\mathbf{F} + \mathbf{F}^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right]\right\}[\Delta\mathbf{u}] = \mathrm{sym}\left((\mathrm{Grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right)$$
5. Use $\delta\mathbf{e} = \delta\boldsymbol{\varepsilon} = \frac{1}{2}\left[(\mathrm{grad}\delta\mathbf{u})^{\mathrm{T}} + \mathrm{grad}\delta\mathbf{u}\right]$ to show that the
$$\begin{aligned}\Delta\delta\mathbf{e} &= \chi_*\left(\partial_{\mathbf{u}}\left(\chi_*^{-1}(\delta\mathbf{e})\right)[\Delta\mathbf{u}]\right) = \chi_*\mathrm{sym}\left((\mathrm{Grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right) \\ &= \mathrm{sym}\left[(\mathrm{grad}\Delta\mathbf{u})^{\mathrm{T}} \cdot \mathrm{grad}\delta\mathbf{u}\right]\end{aligned}$$

3 Stress and the Balance Principles

Three basic laws of physics are discussed in this Chapter:

(1) The Law of Conservation of Mass
(2) The Balance of Linear Momentum
(3) The Balance of Angular Momentum

together with the conservation of mechanical energy and the principle of virtual work, which are different versions of (2).

(2) and (3) involve the concept of stress, which allows one to describe the action of forces in materials.

3.1 Conservation of Mass

3.1.1 Mass and Density

Mass is a non-negative scalar measure of a body's tendency to resist a change in motion.

Consider a small volume element Δv whose mass is Δm. Define the average **density** of this volume element by the ratio

$$\rho_{\text{AVE}} = \frac{\Delta m}{\Delta v} \tag{3.1.1}$$

If p is some point within the volume element, then define the **spatial mass density** at p to be the limiting value of this ratio as the volume shrinks down to the point,

$$\boxed{\rho(\mathbf{x},t) = \lim_{\Delta v \to 0} \frac{\Delta m}{\Delta v}} \qquad \textbf{Spatial Density} \tag{3.1.2}$$

In a real material, the incremental volume element Δv must not actually get too small since then the limit ρ would depend on the atomistic structure of the material; the volume is only allowed to decrease to some minimum value which contains a large number of molecules. The spatial mass density is a representative average obtained by having Δv large compared to the atomic scale, but small compared to a typical length scale of the problem under consideration.

The density, as with displacement, velocity, and other quantities, is defined for *specific particles* of a continuum, and is a continuous function of coordinates and time, $\rho = \rho(\mathbf{x},t)$. However, the mass is not defined this way – one writes for the mass of an infinitesimal volume of material – a **mass element**,

$$dm = \rho(\mathbf{x},t)dv \tag{3.1.3}$$

or, for the mass of a volume v of material at time t,

$$m = \int_v \rho(\mathbf{x},t)dv \tag{3.1.4}$$

3.1.2 Conservation of Mass

The law of conservation of mass states that mass can neither be created nor destroyed.

Consider a collection of matter located somewhere in space. This quantity of matter with well-defined boundaries is termed a **system**. The law of conservation of mass then implies that the mass of this given system remains constant,

$$\boxed{\frac{Dm}{Dt} = 0} \qquad \textbf{Conservation of Mass} \qquad (3.1.5)$$

The volume occupied by the matter may be changing and the density of the matter within the system may be changing, but the mass remains constant.

Considering a differential mass element at position **X** in the reference configuration and at **x** in the current configuration, Eqn. 3.1.5 can be rewritten as

$$dm(\mathbf{X}) = dm(\mathbf{x}, t) \qquad (3.1.6)$$

The conservation of mass equation can be expressed in terms of densities. First, introduce ρ_0, the **reference mass density** (or simply the **density**), defined through

$$\boxed{\rho_0(\mathbf{X}) = \lim_{\Delta V \to 0} \frac{\Delta m}{\Delta V}} \qquad \textbf{Density} \qquad (3.1.7)$$

Note that the density ρ_0 and the spatial mass density ρ are *not* the same quantities[1].

Thus the **local** (or **differential**) **form** of the conservation of mass can be expressed as (see Fig. 3.1.1)

$$dm = \rho_0(\mathbf{X})dV = \rho(\mathbf{x}, t)dv = \text{const} \qquad (3.1.8)$$

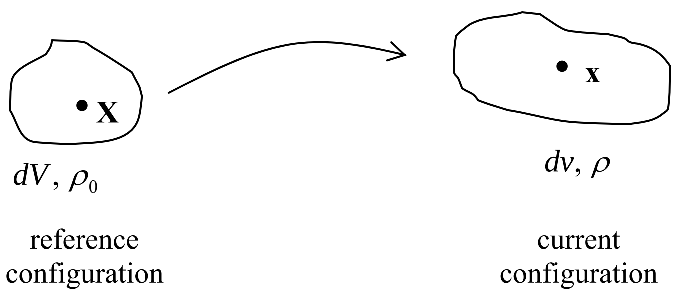


Figure 3.1.1: Conservation of Mass for a deforming mass element

Integration over a finite region of material gives the **global** (or **integral**) **form**,

$$m = \int_V \rho_0(\mathbf{X})dV = \int_v \rho(\mathbf{x}, t)dv = \text{const} \qquad (3.1.9)$$

or

$$\dot{m} = \frac{dm}{dt} = \frac{d}{dt}\int_v \rho(\mathbf{x}, t)dv = 0 \qquad (3.1.10)$$

[1] they not only are functions of different variables, but also have different values; they are not different representations of the same thing, as were, for example, the velocities **v** and **V**. One could introduce a material mass density, $\mathrm{P}(X, t) = \rho(x(X, t), t)$, but such a quantity is not useful in analysis

3.1.3 Control Mass and Control Volume

A **control mass** is a *fixed mass* of material whose volume and density may change, and which may move through space, Fig. 3.1.2. There is no mass transport through the moving surface of the control mass. For such a system, Eqn. 3.1.10 holds.

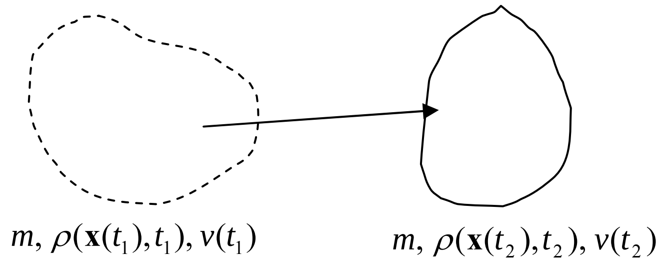


Figure 3.1.2: Control Mass

By definition, the derivative in 3.1.10 is the time derivative of a property (in this case mass) of a collection of material particles as they move through space, and when they instantaneously occupy the volume *v*, Fig. 3.1.3, or

$$\frac{d}{dt}\int_{dv}\rho\, dv = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\int_{v(t+\Delta t)}\rho(\mathbf{x},t+\Delta t)dv - \int_{v(t)}\rho(\mathbf{x},t)dv\right\} = 0 \qquad (3.1.11)$$

Alternatively, one can take the material derivative inside the integral sign:

$$\frac{dm}{dt} = \int_v \frac{d}{dt}\left[\rho(\mathbf{x},t)dv\right] = 0 \qquad (3.1.12)$$

This is now equivalent to the sum of the rates of change of mass of the mass elements occupying the volume *v*.

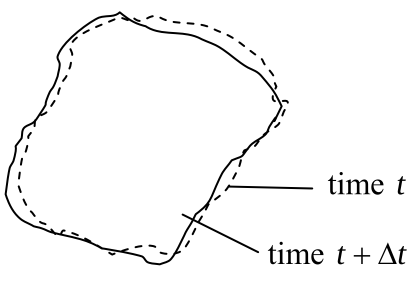


Figure 3.1.3: Control Mass occupying different volumes at different times

A **control volume**, on the other hand, is a *fixed volume* (region) of space through which material may flow, Fig. 3.1.4, and for which *the mass may change*. For such a system, one has

$$\frac{\partial m}{\partial t} = \frac{\partial}{\partial t}\int_v \rho(\mathbf{x},t)dv = \int_v \frac{\partial}{\partial t}[\rho(\mathbf{x},t)]dv \neq 0 \tag{3.1.13}$$

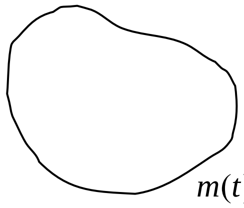


Figure 3.1.4: Control Volume

3.1.4 The Continuity Equation (Spatial Form)

A consequence of the law of conservation of mass is the **continuity equation**, which (in the spatial form) relates the density and velocity of any material particle during motion. This equation can be derived in a number of ways:

Derivation of the Continuity Equation using a Control Volume (Global Form)

The continuity equation can be derived directly by considering a control volume - this is the derivation appropriate to fluid mechanics. Mass inside this fixed volume cannot be created or destroyed, so that the rate of increase of mass in the volume must equal the rate at which mass is flowing into the volume through its bounding surface.

The rate of increase of mass inside the fixed volume v is

$$\frac{\partial m}{\partial t} = \frac{\partial}{\partial t}\int_v \rho(\mathbf{x},t)\,dv = \int_v \frac{\partial \rho}{\partial t}dv \tag{3.1.14}$$

The **mass flux** (rate of flow of mass) out through the surface is given by Eqn. 1.7.9,

$$\int_s \rho\mathbf{v}\cdot\mathbf{n}ds, \qquad \int_s \rho v_i n_i\,ds$$

where **n** is the unit outward normal to the surface and **v** is the velocity. It follows that

$$\int_v \frac{\partial \rho}{\partial t}dv + \int_s \rho\mathbf{v}\cdot\mathbf{n}ds = 0, \qquad \int_v \frac{\partial \rho}{\partial t}dv + \int_s \rho v_i n_i\,ds = 0 \tag{3.1.15}$$

Use of the divergence theorem 1.7.12 leads to

$$\int_v \left[\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho\mathbf{v})\right]dv = 0, \qquad \int_v \left[\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v_i)}{\partial x_i}\right]dv = 0 \tag{3.1.16}$$

leading to the continuity equation,

$$\boxed{\begin{array}{ll} \dfrac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0 & \dfrac{\partial \rho}{\partial t} + \dfrac{\partial (\rho v_i)}{\partial x_i} = 0 \\ \dfrac{d\rho}{dt} + \rho \operatorname{div} \mathbf{v} = 0 & \dfrac{d\rho}{dt} + \rho \dfrac{\partial v_i}{\partial x_i} = 0 \\ \dfrac{\partial \rho}{\partial t} + \operatorname{grad}\rho \cdot \mathbf{v} + \rho \operatorname{div} \mathbf{v} = 0 & \dfrac{\partial \rho}{\partial t} + \dfrac{\partial \rho}{\partial x_i} v_i + \rho \dfrac{\partial v_i}{\partial x_i} = 0 \end{array}} \quad \textbf{Continuity Equation}$$

(3.1.17)

This is (these are) the continuity equation in spatial form. The second and third forms of the equation are obtained by re-writing the local derivative in terms of the material derivative 2.4.7 (see also 1.6.23b).

If the material is incompressible, so the density remains constant in the neighbourhood of a particle as it moves, then the continuity equation reduces to

$$\boxed{\operatorname{div} \mathbf{v} = 0, \quad \frac{\partial v_i}{\partial x_i} = 0} \quad \textbf{Continuity Eqn. for Incompressible Material} \qquad (3.1.18)$$

Derivation of the Continuity Equation using a Control Mass

Here follow two ways to derive the continuity equation using a control mass.

1. Derivation using the Formal Definition

From 3.1.11, adding and subtracting a term:

$$\begin{aligned} \frac{d}{dt} \int_{dv} \rho \, dv = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Bigg\{ & \left[\int_{v(t+\Delta t)} \rho(\mathbf{x}, t + \Delta t) dv - \int_{v(t)} \rho(\mathbf{x}, t + \Delta t) dv \right] \\ & + \left[\int_{v(t)} \rho(\mathbf{x}, t + \Delta t) dv - \int_{v(t)} \rho(\mathbf{x}, t) dv \right] \Bigg\} \end{aligned} \qquad (3.1.19)$$

The terms in the second square bracket correspond to holding the volume v fixed and evidently equals the local rate of change:

$$\frac{d}{dt} \int_{dv} \rho \, dv = \int_{v} \frac{\partial \rho}{\partial t} dv + \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{v(t+\Delta t) - v(t)} \rho(\mathbf{x}, t + \Delta t) dv \qquad (3.1.20)$$

The region $v(t + \Delta t) - v(t)$ is swept out in time Δt. Superimposing the volumes $v(t)$ and $v(t + \Delta t)$, Fig. 3.1.5, it can be seen that a small element Δv of $v(t + \Delta t) - v(t)$ is given by (see the example associated with Fig. 1.7.7)

$$\Delta v = \Delta t \mathbf{v} \cdot \mathbf{n} \Delta s \qquad (3.1.21)$$

where *s* is the surface. Thus

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{v(t+\Delta t)-v(t)} \rho(\mathbf{x}, t+\Delta t) dv = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_s \Delta t \rho(\mathbf{x}, t+\Delta t) \mathbf{v} \cdot \mathbf{n} ds = \int_s \rho(\mathbf{x}, t) \mathbf{v} \cdot \mathbf{n} ds \quad (3.1.22)$$

and 3.1.15 is again obtained, from which the continuity equation results from use of the divergence theorem.

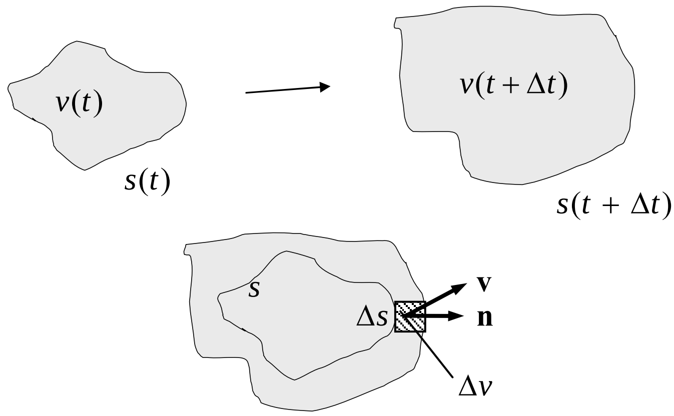


Figure 3.1.5: Evaluation of Eqn. 3.1.22

2. Derivation by Converting to Mass Elements

This derivation requires the kinematic relation for the material time derivative of a volume element, 2.5.23: $d(dv)/dt = \mathrm{div}\mathbf{v}\, dv$. One has

$$\frac{dm}{dt} = \frac{d}{dt} \int_v \rho(\mathbf{x}, t) dv = \int_v \frac{d}{dt}(\rho dv) = \int_v \left(\dot{\rho} dv + \rho \dot{\overline{dv}} \right) = \int_v (\dot{\rho} + \mathrm{div}\mathbf{v} \rho) dv \equiv 0 \qquad (3.1.23)$$

The continuity equation then follows, since this must hold for any arbitrary region of the volume *v*.

Derivation of the Continuity Equation using a Control Volume (Local Form)

The continuity equation can also be derived using a differential control volume element. This calculation is similar to that given in §1.6.6, with the velocity $\mathbf{v}$ replaced by $\rho\mathbf{v}$.

3.1.5 The Continuity Equation (Material Form)

From 3.1.9, and using 2.2.53, $dv = JdV$,

$$\int_V [\rho_0(\mathbf{X}) - \rho(\boldsymbol{\chi}(\mathbf{X}, t), t) J(\mathbf{X}, t)] dV = 0 \qquad (3.1.24)$$

Since V is an arbitrary region, the integrand must vanish everywhere, so that

$$\boxed{\rho_0(\mathbf{X}) = \rho(\chi(\mathbf{X},t),t)J(\mathbf{X},t)} \quad \textbf{Continuity Equation (Material Form)} \qquad (3.1.25)$$

This is known as the continuity (mass) equation in the material description. Since $\dot{\rho}_0 = 0$, the rate form of this equation is simply

$$\frac{d}{dt}(\rho J) = 0 \qquad (3.1.26)$$

The material form of the continuity equation, $\rho_0 = \rho J$, is an algebraic equation, unlike the partial differential equation in the spatial form. However, the two must be equivalent, and indeed the spatial form can be derived directly from this material form: using 2.5.20, $dJ/dt = J\mathrm{div}\mathbf{v}$,

$$\begin{aligned} \frac{d}{dt}(\rho J) &= \dot{\rho}J + \rho\dot{J} \\ &= J\left(\dot{\rho} + \rho\,\mathrm{div}\mathbf{v}\right) \end{aligned} \qquad (3.1.27)$$

This is zero, and $J > 0$, and the spatial continuity equation follows.

Example (of Conservation of Mass)

Consider a bar of material of length l_0, with density in the undeformed configuration ρ_0 and spatial mass density $\rho(x,t)$, undergoing the 1-D motion $\mathbf{X} = \mathbf{x}/(1+At)$, $\mathbf{x} = \mathbf{X} + At\mathbf{X}$. The volume ratio (taking unit cross-sectional area) is $J = 1 + At$. The continuity equation in the material form 3.1.25 specifies that

$$\rho_0 = \rho(1+At)$$

Suppose now that

$$\rho_o(\mathbf{X}) = \frac{2m}{l_0^2}\mathbf{X}$$

so that the total mass of the bar is $\int_0^{l_o} \rho_0(\mathbf{X})d\mathbf{X} = m$. It follows that the spatial mass density is

$$\rho = \frac{\rho_0}{(1+At)} = \frac{2m}{l_0^2}\frac{\mathbf{X}}{1+At} = \frac{2m}{l_0^2}\frac{\mathbf{x}}{(1+At)^2}$$

Evaluating the total mass of the bar at time t leads to

$$\int_0^{l_o(1+At)} \rho(\mathbf{x},t)\,d\mathbf{x} = \frac{2m}{l_0^2}\frac{1}{(1+At)^2}\int_0^{l_o(1+At)} \mathbf{x}\,d\mathbf{x}$$

which is again *m*, as required.

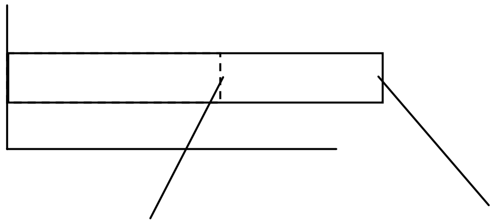


Figure 3.1.6: a stretching bar

The density could have been derived from the equation of continuity in the spatial form: since the velocity is

$$\mathbf{V}(\mathbf{X},t) = \frac{d\mathbf{x}(\mathbf{X},t)}{dt} = A\mathbf{X}, \qquad \mathbf{v}(\mathbf{x},t) = \mathbf{V}(\chi^{-1}(\mathbf{x},t),t) = \frac{A\mathbf{x}}{1+At}$$

one has

$$\frac{\partial \rho}{\partial t} + \mathbf{v}\frac{\partial \rho}{\partial \mathbf{x}} + \rho\frac{\partial v}{\partial \mathbf{x}} = \frac{\partial \rho}{\partial t} + \frac{A\mathbf{x}}{1+At}\frac{\partial \rho}{\partial \mathbf{x}} + \rho\frac{A}{1+At} = 0$$

Without attempting to solve this first order partial differential equation, it can be seen by substitution that the value for ρ obtained previously satisfies the equation.

■

3.1.6 Material Derivatives of Integrals

Reynold's Transport Theorem

In the above, the material derivative of the total mass carried by a control mass,

$$\frac{d}{dt}\int_v \rho(\mathbf{x},t)dv,$$

was considered. It is quite often that one needs to evaluate material time derivatives of similar volume (and line and surface) integrals, involving other properties, for example momentum or energy. Thus, suppose that $\mathbf{A}(\mathbf{x},t)$ is the distribution of some property (per unit volume) throughout a volume *v* (**A** is taken to be a second order tensor, but what follows applies also to vectors and scalars). Then the rate of change of the total amount of the property carried by the mass system is

$$\frac{d}{dt}\int_v \mathbf{A}(\mathbf{x},t)dv$$

Again, this integral can be evaluated in a number of ways. For example, one could evaluate it using the formal definition of the material derivative, as done above for $\mathbf{A} = \rho$. Alternatively, one can evaluate it using the relation 2.5.23, $d(dv)/dt = \operatorname{div}\mathbf{v}\,dv$, through

$$\frac{d}{dt}\int_v \mathbf{A}(\mathbf{x},t)dv = \int_v \frac{d}{dt}\left[\mathbf{A}(\mathbf{x},t)dv\right] = \int_v \left[\dot{\mathbf{A}}dv + \mathbf{A}\,\dot{\overline{dv}}\right] = \int_v \left[\dot{\mathbf{A}} + \operatorname{div}\mathbf{v}\mathbf{A}\right]dv \qquad (3.1.28)$$

Thus one arrives at **Reynold's transport theorem**

$$\frac{d}{dt}\int_v \mathbf{A}(\mathbf{x},t)dv = \begin{cases} \displaystyle\int_v \left[\frac{d\mathbf{A}}{dt} + \operatorname{div}\mathbf{v}\mathbf{A}\right]dv & \displaystyle\int_v \left[\frac{dA_{ij}}{dt} + \frac{\partial v_k}{\partial x_k}A_{ij}\right]dv \\ \displaystyle\int_v \left[\frac{\partial\mathbf{A}}{\partial t} + \operatorname{grad}\mathbf{A}\cdot\mathbf{v} + \operatorname{div}\mathbf{v}\mathbf{A}\right]dv & \displaystyle\int_v \left[\frac{\partial A_{ij}}{\partial t} + \frac{\partial A_{ij}}{\partial x_k}v_k + \frac{\partial v_k}{\partial x_k}A_{ij}\right]dv \\ \displaystyle\int_v \left[\frac{\partial\mathbf{A}}{\partial t} + \operatorname{div}(\mathbf{A}\otimes\mathbf{v})\right]dv & \displaystyle\int_v \left[\frac{\partial A_{ij}}{\partial t} + \frac{\partial(A_{ij}v_k)}{\partial x_k}\right]dv \\ \displaystyle\int_v \frac{\partial\mathbf{A}}{\partial t}dv + \int_s \mathbf{A}(\mathbf{v}\cdot\mathbf{n})ds & \displaystyle\int_v \frac{\partial A_{ij}}{\partial t}dv + \int_s A_{ij}v_k n_k ds \end{cases}$$

Reynold's Transport Theorem (3.1.29)

The index notation is shown for the case when **A** is a second order tensor. In the last of these forms[2] (obtained by application of the divergence theorem), the first term represents the amount (of **A**) created within the volume *v* whereas the second term (the flux term) represents the (volume) rate of flow of the property through the surface. In the last three versions, Reynold's transport theorem gives the material derivative of the moving control mass in terms of the derivative of the instantaneous fixed volume in space (the first term).

Of course when $\mathbf{A} = \rho$, the continuity equation is recovered.

Another way to derive this result is to first convert to the reference configuration, so that integration and differentiation commute (since dV is independent of time):

$$\begin{aligned} \frac{d}{dt}\int_v \mathbf{A}(\mathbf{x},t)dv &= \frac{d}{dt}\int_V \mathbf{A}(\mathbf{X},t)JdV = \int_V \frac{d}{dt}\left(\mathbf{A}(\mathbf{X},t)J\right)dV \\ &= \int_V \left(\dot{\mathbf{A}}J + \mathbf{A}\dot{J}\right)dV = \int_V \left(\dot{\mathbf{A}} + \operatorname{div}\mathbf{v}\mathbf{A}\right)JdV \\ &= \int_v \left(\dot{\mathbf{A}}(\mathbf{x},t) + \operatorname{div}\mathbf{v}\mathbf{A}(\mathbf{x},t)\right)dv \end{aligned} \qquad (3.1.30)$$

[2] also known as the **Leibniz formula**

Reynold's Transport Theorem for Specific Properties

A property that is given per unit mass is called a **specific property**. For example, specific heat is the heat per unit mass. Consider then a property **B**, a scalar, vector or tensor, which is defined per unit mass through a volume. Then the rate of change of the total amount of the property carried by the mass system is simply

$$\frac{d}{dt}\int_v \rho\mathbf{B}(\mathbf{x},t)dv = \int_v \frac{d}{dt}\left[\mathbf{B}\rho dv\right] = \int_v \frac{d}{dt}\left[\mathbf{B}dm\right] = \int_v \frac{d\mathbf{B}}{dt}dm = \int_v \rho\frac{d\mathbf{B}}{dt}dv \qquad (3.1.31)$$

Material Derivatives of Line and Surface Integrals

Material derivatives of line and surface integrals can also be evaluated. From 2.5.8, $d(d\mathbf{x})/dt = \mathbf{l}d\mathbf{x}$,

$$\frac{d}{dt}\int \mathbf{A}(\mathbf{x},t)d\mathbf{x} = \int\left[\dot{\mathbf{A}} + \mathbf{A}\mathbf{l}\right]d\mathbf{x} \qquad (3.1.32)$$

and, using 2.5.22, $d\left(\hat{\mathbf{n}}ds\right)/dt = \left(\mathrm{div}\mathbf{v} - \mathbf{l}^{\mathrm{T}}\right)\hat{\mathbf{n}}ds$,

$$\frac{d}{dt}\int_s \mathbf{A}(\mathbf{x},t)\hat{\mathbf{n}}ds = \int_s\left[\dot{\mathbf{A}} + \mathbf{A}\left(\mathrm{div}\mathbf{v} - \mathbf{l}^{\mathrm{T}}\right)\right]\hat{\mathbf{n}}ds \qquad (3.1.33)$$

3.1.7 Problems

1. A motion is given by the equations

$$x_1 = X_1 + 3X_2 t, \quad x_2 = -X_1 t^2 + X_2(t+1), \quad x_3 = X_3$$

(a) Calculate the spatial mass density ρ in terms of the density ρ_0

(b) Derive a first order ordinary differential equation for the density ρ (in terms of **x** and t only) assuming that it is independent of position **x**

3.2 The Momentum Principles

In Parts I and II, the basic dynamics principles used were Newton's Laws, and these are equivalent to force equilibrium and moment equilibrium. For example, they were used to derive the stress transformation equations in Part I, §3.4 and the Equations of Motion in Part II, §1.1. Newton's laws there were applied to differential material elements.

An alternative but completely equivalent set of dynamics laws are **Euler's Laws**; these are more appropriate for finite-sized collections of moving particles, and can be used to express the force and moment equilibrium in terms of integrals. Euler's Laws are also called the **Momentum Principles**: the **principle of linear momentum** (Euler's first law) and the **principle of angular momentum** (Euler's second law).

3.2.1 The Principle of Linear Momentum

Momentum is a measure of the tendency of an object to keep moving once it is set in motion. Consider first the particle of rigid body dynamics: the (linear) momentum $\mathbf{p}$ is defined to be its mass times velocity, $\mathbf{p} = m\mathbf{v}$. The rate of change of momentum $\dot{\mathbf{p}}$ is

$$\frac{d\mathbf{p}}{dt} = \frac{d(m\mathbf{v})}{dt} = m\frac{d\mathbf{v}}{dt} = m\mathbf{a} \tag{3.2.1}$$

and use has been made of the fact that $dm/dt = 0$. Thus Newton's second law, $\mathbf{F} = m\mathbf{a}$, can be rewritten as

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}) \tag{3.2.2}$$

This equation, formulated by Euler, states that *the rate of change of momentum is equal to the applied force*. It is called the **principle of linear momentum**, or **balance of linear momentum**. If there are no forces applied to a system, the total momentum of the system remains constant; the law in this case is known as the **law of conservation of (linear) momentum**.

Eqn. 3.2.2 as applied to a particle can be generalized to the mechanics of a continuum in one of two ways. One could consider a differential element of material, of mass dm and velocity $\mathbf{v}$. Alternatively, one can consider a finite portion of material, a control mass in the current configuration with spatial mass density $\rho(\mathbf{x},t)$ and spatial velocity field $\mathbf{v}(\mathbf{x},t)$. The total linear momentum of this mass of material is

$$\boxed{\mathbf{L}(t) = \int_v \rho(\mathbf{x},t)\mathbf{v}(\mathbf{x},t)dv} \quad \textbf{Linear Momentum} \tag{3.2.3}$$

The principle of linear momentum states that

$$\dot{\mathbf{L}}(t) = \frac{d}{dt}\int_v \rho(\mathbf{x},t)\mathbf{v}(\mathbf{x},t)dv = \mathbf{F}(t) \tag{3.2.4}$$

where $\mathbf{F}(t)$ is the resultant of the forces acting on the portion of material.

Note that the volume over which the integration in Eqn. 3.2.4 takes place is not fixed; the integral is taken over a *fixed portion of material particles*, and the space occupied by this matter may change over time.

By virtue of the Transport theorem relation 3.1.31, this can be written as

$$\dot{\mathbf{L}}(t) = \int_v \rho(\mathbf{x},t)\dot{\mathbf{v}}(\mathbf{x},t)dv = \mathbf{F}(t) \tag{3.2.5}$$

The resultant force acting on a body is due to the surface tractions **t** acting over surface elements and body forces **b** acting on volume elements, Fig. 3.2.1:

$$\boxed{\mathbf{F}(t) = \int_s \mathbf{t}ds + \int_v \mathbf{b}dv, \qquad F_i = \int_s t_i\, ds + \int_v b_i\, dv} \qquad \textbf{Resultant Force} \tag{3.2.6}$$

and so the principle of linear momentum can be expressed as

$$\boxed{\int_s \mathbf{t}ds + \int_v \mathbf{b}dv = \int_v \rho\dot{\mathbf{v}}dv} \qquad \textbf{Principle of Linear Momentum} \tag{3.2.7}$$

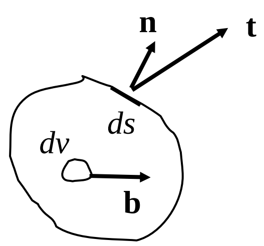


Figure 3.2.1: surface and body forces acting on a finite volume of material

The principle of linear momentum, Eqns. 3.2.7, will be used to prove Cauchy's Lemma and Cauchy's Law in the next section and, in §3.6, to derive the Equations of Motion.

3.2.2 The Principle of Angular Momentum

Considering again the mechanics of a single particle: the **angular momentum** is the moment of momentum about an axis, in other words, it is the product of the linear momentum of the particle and the perpendicular distance from the axis of its line of action. In the notation of Fig. 3.2.2, the angular momentum **h** is

$$\mathbf{h} = \mathbf{r} \times m\mathbf{v} \tag{3.2.8}$$

which is the vector with magnitude $d \times m|\mathbf{v}|$ and perpendicular to the plane shown.

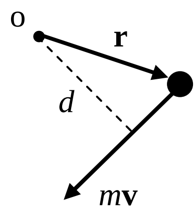


Figure 3.2.2: surface and body forces acting on a finite volume of material

Consider now a collection of particles. The **principle of angular momentum** states that the resultant moment of the external forces acting on the system of particles, $\mathbf{M}$, equals the rate of change of the total angular momentum of the particles:

$$\mathbf{M} = \mathbf{r} \times \mathbf{F} = \frac{d\mathbf{h}}{dt} \tag{3.2.9}$$

Generalising to a continuum, the angular momentum is

$$\boxed{\mathbf{H} = \int_v \mathbf{r} \times \rho\mathbf{v}dv} \qquad \textbf{Angular Momentum} \tag{3.2.10}$$

and the principle of angular momentum is

$$\boxed{\begin{gathered} \int_s \mathbf{r} \times \mathbf{t}^{(\mathbf{n})} ds + \int_v \mathbf{r} \times \mathbf{b} dv = \frac{d}{dt} \int_v \mathbf{r} \times \rho\mathbf{v} dv \\ \int_s \varepsilon_{ijk} x_j t_k^{(\mathbf{n})} ds + \int_v \varepsilon_{ijk} x_j b_k \, dv = \frac{d}{dt} \int_V \varepsilon_{ijk} x_j \rho v_k \, dv \end{gathered}} \qquad \textbf{Principle of Angular Momentum}$$

$$\tag{3.2.11}$$

The principle of angular momentum, 3.2.11, will be used, in §3.6, to deduce the symmetry of the Cauchy stress.

3.3 The Cauchy Stress Tensor

3.3.1 The Traction Vector

The **traction vector** was introduced in Part I, §3.3. To recall, it is the limiting value of the ratio of force over area; for Force ΔF acting on a surface element of area ΔS, it is

$$\mathbf{t}^{(\mathbf{n})} = \lim_{\Delta S \to 0} \frac{\Delta F}{\Delta S} \tag{3.3.1}$$

and $\mathbf{n}$ denotes the normal to the surface element. An infinite number of traction vectors act at a point, each acting on different surfaces through the point, defined by different normals.

3.3.2 Cauchy's Lemma

Cauchy's lemma states that traction vectors acting on opposite sides of a surface are equal and opposite[1]. This can be expressed in vector form:

$$\boxed{\mathbf{t}^{(\mathbf{n})} = -\mathbf{t}^{(-\mathbf{n})}} \quad \textbf{Cauchy's Lemma} \tag{3.3.2}$$

This can be proved by applying the principle of linear momentum to a collection of particles of mass Δm instantaneously occupying a small box with parallel surfaces of area Δs, thickness δ and volume $\Delta v = \delta \Delta s$, Fig. 3.3.1. The resultant *surface* force acting on this matter is $\mathbf{t}^{(\mathbf{n})}\Delta s + \mathbf{t}^{(-\mathbf{n})}\Delta s$.

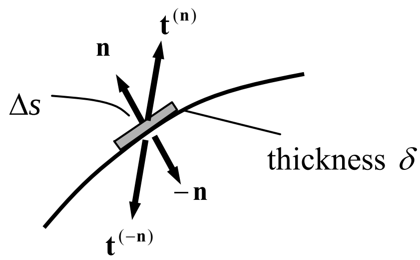


Figure 3.3.1: traction acting on a small portion of material particles

The total linear momentum of the matter is $\int_{\Delta V} \rho \mathbf{v} dv = \int_{\Delta m} \mathbf{v} dm$. By the mean value theorem (see Appendix A to Chapter 1, §1.B.1), this equals $\bar{\mathbf{v}}\Delta m$, where $\bar{\mathbf{v}}$ is the velocity at some interior point. Similarly, the body force acting on the matter is $\int_{\Delta V} \mathbf{b} dv = \bar{\mathbf{b}}\Delta v$, where $\bar{\mathbf{b}}$ is the body force (per unit volume) acting at some interior point. The total mass

[1] this is equivalent to Newton's (third) law of action and reaction – it seems like a lot of work to prove this seemingly obvious result but, to be consistent, it is supposed that the only fundamental dynamic laws available here are the principles of linear and angular momentum, and not any of Newton's laws

can also be written as $\Delta m = \int_{\Delta V} \rho dv = \bar{\rho}\Delta v$. From the principle of linear momentum, Eqn. 3.2.7, and since Δm does not change with time,

$$\mathbf{t}^{(\mathbf{n})}\Delta s + \mathbf{t}^{(-\mathbf{n})}\Delta s + \overline{\mathbf{b}}\Delta v = \frac{d}{dt}\left[\overline{\mathbf{v}}\Delta m\right] = \Delta m \frac{d\overline{\mathbf{v}}}{dt} = \bar{\rho}\Delta v \frac{d\overline{\mathbf{v}}}{dt} = \bar{\rho}\delta\Delta s \frac{d\overline{\mathbf{v}}}{dt} \tag{3.3.3}$$

Dividing through by Δs and taking the limit as $\delta \to 0$, one finds that $\mathbf{t}^{(\mathbf{n})} = -\mathbf{t}^{(-\mathbf{n})}$. Note that the values of $\mathbf{t}^{(\mathbf{n})}$, $\mathbf{t}^{(-\mathbf{n})}$ acting on the box with finite thickness are not the same as the final values, but approach the final values *at* the surface as $\delta \to 0$.

3.3.3 Stress

In Part I, the components of the traction vector were called stress components, and it was illustrated how there were nine stress components associated with each material particle. Here, the stress is defined more formally,

Cauchy's Law

Cauchy's Law states that there exists a **Cauchy stress tensor $\boldsymbol{\sigma}$** which maps the normal to a surface to the traction vector acting on that surface, according to

$$\boxed{\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}, \qquad t_i = \sigma_{ij} n_j} \qquad \textbf{Cauchy's Law} \tag{3.3.4}$$

or, in full,

$$\begin{aligned} t_1 &= \sigma_{11} n_1 + \sigma_{12} n_2 + \sigma_{13} n_3 \\ t_2 &= \sigma_{21} n_1 + \sigma_{22} n_2 + \sigma_{23} n_3 \\ t_3 &= \sigma_{31} n_1 + \sigma_{32} n_2 + \sigma_{33} n_3 \end{aligned} \tag{3.3.5}$$

Note:

- many authors define the stress tensor as $\mathbf{t} = \mathbf{n}\boldsymbol{\sigma}$. This amounts to the definition used here since, as mentioned in Part I, and as will be (re-)proved below, the stress tensor is symmetric, $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathrm{T}}$, $\sigma_{ij} = \sigma_{ji}$
- the Cauchy stress refers to the *current* configuration, that is, it is a measure of force per unit area acting on a surface in the current configuration.

Stress Components

Taking Cauchy's law to be true (it is proved below), the components of the stress tensor with respect to a Cartesian coordinate system are, from 1.9.4 and 3.3.4,

$$\sigma_{ij} = \mathbf{e}_i \boldsymbol{\sigma} \mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{t}^{(\mathbf{e}_j)} \tag{3.3.6}$$

which is the *i*th component of the traction vector acting on a surface with normal $\mathbf{e}_j$. Note that this definition is inconsistent with that given in Part I, §3.2 – there, the first subscript denoted the direction of the normal – but, again, the two definitions are equivalent because of the symmetry of the stress tensor.

The three traction vectors acting on the surface elements whose outward normals point in the directions of the three base vectors $\mathbf{e}_j$ are

$$\mathbf{t}^{(\mathbf{e}_j)} = \boldsymbol{\sigma}\mathbf{e}_j, \qquad \begin{aligned} \mathbf{t}^{(\mathbf{e}_1)} &= \sigma_{11}\mathbf{e}_1 + \sigma_{21}\mathbf{e}_2 + \sigma_{31}\mathbf{e}_3 \\ \mathbf{t}^{(\mathbf{e}_2)} &= \sigma_{12}\mathbf{e}_1 + \sigma_{22}\mathbf{e}_2 + \sigma_{32}\mathbf{e}_3 \\ \mathbf{t}^{(\mathbf{e}_3)} &= \sigma_{13}\mathbf{e}_1 + \sigma_{23}\mathbf{e}_2 + \sigma_{33}\mathbf{e}_3 \end{aligned} \tag{3.3.7}$$

Eqns. 3.3.6-7 are illustrated in Fig. 3.3.2.

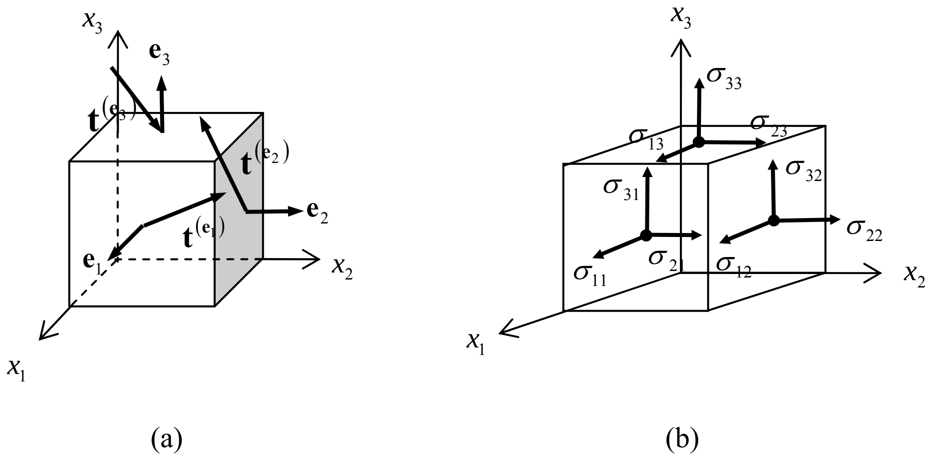


Figure 3.3.2: traction acting on surfaces with normals in the coordinate directions; (a) traction vectors, (b) stress components

Proof of Cauchy's Law

The proof of Cauchy's law essentially follows the same method as used in the proof of Cauchy's lemma.

Consider a small tetrahedral free-body, with vertex at the origin, Fig. 3.3.3. It is required to determine the traction $\mathbf{t}$ in terms of the nine stress components (which are all shown positive in the diagram).

Let the area of the base of the tetrahedran, with normal $\mathbf{n}$, be Δs. The area ds_1 is then $\Delta s \cos\alpha$, where α is the angle between the planes, as shown in Fig. 3.3.3b; this angle is the same as that between the vectors $\mathbf{n}$ and $\mathbf{e}_1$, so $\Delta s_1 = (\mathbf{n} \cdot \mathbf{e}_1)\Delta s = n_1 \Delta s$, and similarly for the other surfaces: $\Delta s_2 = n_2 \Delta s$ and $\Delta s_3 = n_3 \Delta s$.

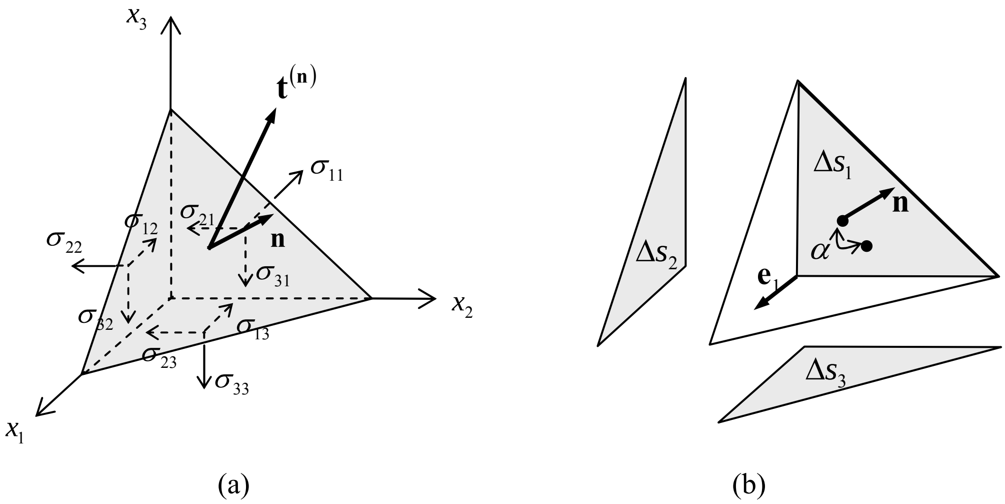


Figure 3.3.3: free body diagram of a tetrahedral portion of material; (a) traction acting on the material, (b) relationship between surface areas and normal components

The resultant surface force on the body, acting in the x_1 direction, is

$$t_1\Delta s - \sigma_{11}n_1\Delta s - \sigma_{12}n_2\Delta s - \sigma_{13}n_3\Delta s$$

Again, the momentum is $\overline{\mathbf{v}}\Delta M$, the body force is $\overline{\mathbf{b}}\Delta v$ and the mass is $\Delta m = \overline{\rho}\Delta v = \overline{\rho}(h/3)\Delta s$, where h is the perpendicular distance from the origin (vertex) to the base. The principle of linear momentum then states that

$$t_1\Delta s - \sigma_{11}n_1\Delta s - \sigma_{12}n_2\Delta s - \sigma_{13}n_3\Delta s + \overline{b}_1(h/3)\Delta s = \overline{\rho}(h/3)\Delta s\frac{d\overline{v}_1}{dt}$$

Again, the values of the traction and stress components on the faces will in general vary over the faces, so the values used in this equation are average values over the faces.

Dividing through by Δs, and taking the limit as $h \to 0$, one finds that

$$t_1 = \sigma_{11}n_1 + \sigma_{12}n_2 + \sigma_{13}n_3$$

and now these quantities, t_1, σ_{11}, σ_{12}, σ_{13}, are the values *at* the origin. The equations for the other two traction components can be derived in a similar way.

Normal and Shear Stress

The stress acting normal to a surface is given by

$$\sigma_N = \mathbf{n}\cdot\mathbf{t}^{(\mathbf{n})} \tag{3.3.8}$$

The shear stress acting on the surface can then be obtained from

$$\sigma_S = \sqrt{\left|\mathbf{t}^{(\mathbf{n})}\right|^2 - \sigma_N^2} \tag{3.3.9}$$

Example

The state of stress at a point is given in the matrix form

$$[\sigma_{ij}] = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -2 \\ 3 & -2 & 1 \end{bmatrix}$$

Determine

(a) the traction vector acting on a plane through the point whose unit normal is $\hat{\mathbf{n}} = (1/3)\hat{\mathbf{e}}_1 + (2/3)\hat{\mathbf{e}}_2 - (2/3)\hat{\mathbf{e}}_3$

(b) the component of this traction acting perpendicular to the plane

(c) the shear component of traction.

<u>Solution</u>

(a) The traction is

$$\begin{bmatrix} t_1^{(\hat{\mathbf{n}})} \\ t_2^{(\hat{\mathbf{n}})} \\ t_3^{(\hat{\mathbf{n}})} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -2 \\ 3 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -2 \\ 9 \\ -3 \end{bmatrix}$$

or $\mathbf{t}^{(\hat{\mathbf{n}})} = (-2/3)\hat{\mathbf{e}}_1 + 3\hat{\mathbf{e}}_2 - \hat{\mathbf{e}}_3$.

(b) The component normal to the plane is the projection of $\mathbf{t}^{(\hat{\mathbf{n}})}$ in the direction of $\hat{\mathbf{n}}$, i.e.

$$\sigma_N = \mathbf{t}^{(\hat{\mathbf{n}})} \cdot \hat{\mathbf{n}} = (-2/3)(1/3) + 3(2/3) + (2/3) = 22/9 \approx 2.4.$$

(c) The shearing component of traction is

$$\begin{aligned} \sigma_S &= \mathbf{t}^{(\hat{\mathbf{n}})} - (22/9)\hat{\mathbf{n}} \\ &= \left[\left[(-2/3) - (22/27)\right]\hat{\mathbf{e}}_1 + \left[3 - (44/27)\right]\hat{\mathbf{e}}_2 + \left[-1 + (44/27)\right]\hat{\mathbf{e}}_3\right] \\ &= \left[(-40/27)\hat{\mathbf{e}}_1 + (37/27)\hat{\mathbf{e}}_2 + (17/27)\hat{\mathbf{e}}_3\right] \end{aligned}$$

i.e. of magnitude $\sqrt{(-40/27)^2 + (37/27)^2 + (17/27)^2} \approx 2.1$, which equals $\sqrt{\left|\mathbf{t}^{(\hat{\mathbf{n}})}\right|^2 - \sigma_N^2}$.

■

3.4 Properties of the Stress Tensor

3.4.1 Stress Transformation

Let the components of the Cauchy stress tensor in a coordinate system with base vectors $\mathbf{e}_i$ be σ_{ij}. The components in a second coordinate system with base vectors $\mathbf{e}'_j$, σ'_{ij}, are given by the tensor transformation rule 1.10.5:

$$\sigma'_{ij} = Q_{pi} Q_{qj} \sigma_{pq} \tag{3.4.1}$$

where Q_{ij} are the direction cosines, $Q_{ij} = \mathbf{e}_i \cdot \mathbf{e}'_j$.

Isotropic State of Stress

Suppose the state of stress in a body is

$$[\boldsymbol{\sigma}] = \begin{bmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_0 & 0 \\ 0 & 0 & \sigma_0 \end{bmatrix}$$

One finds that the application of the tensor transformation rule yields the very same components no matter what the coordinate system. This is termed an **isotropic** state of stress, or a **spherical** state of stress (see §1.13.3). One example of isotropic stress is the stress arising in fluid at rest, which cannot support shear stress, in which case

$$\boldsymbol{\sigma} = -p\mathbf{I} \tag{3.4.2}$$

where the scalar p is the fluid **hydrostatic pressure**. For this reason, an isotropic state of stress is also referred to as a **hydrostatic** state of stress.

A note on the Transformation Formula

Using the vector transformation rule 1.5.5, the traction and normal transform according to $[\mathbf{t}'] = [\mathbf{Q}^{\mathrm{T}}][\mathbf{t}]$, $[\mathbf{n}'] = [\mathbf{Q}^{\mathrm{T}}][\mathbf{n}]$. Also, Cauchy's law transforms according to $[\mathbf{t}'] = [\boldsymbol{\sigma}'][\mathbf{n}']$ which can be written as $[\mathbf{Q}^{\mathrm{T}}][\mathbf{t}] = [\boldsymbol{\sigma}'][\mathbf{Q}^{\mathrm{T}}][\mathbf{n}]$, so that, pre-multiplying by $[\mathbf{Q}]$, and since $[\mathbf{Q}]$ is orthogonal, $[\mathbf{t}] = \{[\mathbf{Q}][\boldsymbol{\sigma}'][\mathbf{Q}^{\mathrm{T}}]\}[\mathbf{n}]$, so $[\boldsymbol{\sigma}] = [\mathbf{Q}][\boldsymbol{\sigma}'][\mathbf{Q}^{\mathrm{T}}]$, which is the inverse tensor transformation rule 1.13.6a, showing the internal consistency of the theory.

In Part I, Newton's law was applied to a material element to derive the two-dimensional stress transformation equations, Eqn. 3.4.7 of Part I. Cauchy's law was proved in a similar way, using the principle of momentum. In fact, Cauchy's law and the stress transformation equations are equivalent. Given the stress components in one coordinate system, the stress transformation equations give the components in a new coordinate system; particularising this, they give the stress components, and thus the traction vector, acting on new surfaces, oriented in some way with respect to the original axes, which is what Cauchy's law does.

3.4.2 Principal Stresses

Since the stress $\boldsymbol{\sigma}$ is a symmetric tensor, it has three real eigenvalues $\sigma_1, \sigma_2, \sigma_3$, called **principal stresses**, and three corresponding orthonormal eigenvectors called **principal directions**. The eigenvalue problem can be written as

$$\mathbf{t}^{(\mathbf{n})} = \boldsymbol{\sigma}\mathbf{n} = \sigma\mathbf{n} \tag{3.4.3}$$

where $\mathbf{n}$ is a principal direction and σ is a scalar principal stress. Since the traction vector is a multiple of the unit normal, σ is a normal stress component. Thus a principal stress is a stress which acts on a plane of zero shear stress, Fig. 3.4.1.

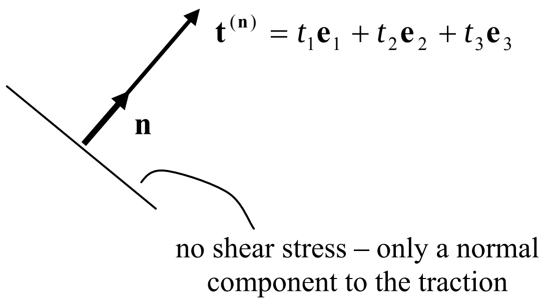


Figure 3.4.1: traction acting on a plane of zero shear stress

The principal stresses are the roots of the characteristic equation 1.11.5,

$$\sigma^3 - \mathrm{I}_1\sigma^2 + \mathrm{I}_2\sigma - \mathrm{I}_3 = 0 \tag{3.4.4}$$

where, Eqn. 1.11.6-7, 1.11.17,

$$\begin{aligned}
I_1 &= \mathrm{tr}\boldsymbol{\sigma} \\
&= \sigma_{11} + \sigma_{22} + \sigma_{33} \\
&= \sigma_1 + \sigma_2 + \sigma_3 \\
I_2 &= \tfrac{1}{2}\left[(\mathrm{tr}\boldsymbol{\sigma})^2 - \mathrm{tr}\boldsymbol{\sigma}^2\right] \\
&= \sigma_{11}\sigma_{22} + \sigma_{22}\sigma_{33} + \sigma_{33}\sigma_{11} - \sigma_{12}^2 - \sigma_{23}^2 - \sigma_{31}^2 \\
&= \sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1 \\
I_3 &= \tfrac{1}{3}\left[\mathrm{tr}\boldsymbol{\sigma}^3 - \tfrac{3}{2}\mathrm{tr}\boldsymbol{\sigma}\mathrm{tr}\boldsymbol{\sigma}^2 + \tfrac{1}{2}(\mathrm{tr}\boldsymbol{\sigma})^3\right] \\
&= \det\boldsymbol{\sigma} \\
&= \sigma_{11}\sigma_{22}\sigma_{33} - \sigma_{11}\sigma_{23}^2 - \sigma_{22}\sigma_{31}^2 - \sigma_{33}\sigma_{12}^2 + 2\sigma_{12}\sigma_{23}\sigma_{32} \\
&= \sigma_1\sigma_2\sigma_3
\end{aligned} \tag{3.4.5}$$

The principal stresses and principal directions are properties of the stress tensor, and do not depend on the particular axes chosen to describe the state of stress., and the **stress invariants** I_1, I_2, I_3 are invariant under coordinate transformation. *c.f.* §1.11.1.

If one chooses a coordinate system to coincide with the three eigenvectors, one has the spectral decomposition 1.11.11 and the stress matrix takes the simple form 1.11.12,

$$\boldsymbol{\sigma} = \sum_{i=1}^{3} \sigma_i \hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i, \qquad [\boldsymbol{\sigma}] = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \tag{3.4.6}$$

Note that when two of the principal stresses are equal, one of the principal directions will be unique, but the other two will be arbitrary – one can choose any two principal directions in the plane perpendicular to the uniquely determined direction, so that the three form an orthonormal set. This stress state is called **axi-symmetric**. When all three principal stresses are equal, one has an isotropic state of stress, and all directions are principal directions.

3.4.3 Maximum Stresses

Directly from §1.11.3, the three principal stresses include the maximum and minimum normal stress components acting at a point. This result is re-derived here, together with results for the maximum shear stress

Normal Stresses

Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be unit vectors *in the principal directions* and consider an arbitrary unit normal vector $\mathbf{n} = n_1\mathbf{e}_1 + n_2\mathbf{e}_2 + n_3\mathbf{e}_3$, Fig. 3.4.2. From 3.3.8 and Cauchy's law, the normal stress acting on the plane with normal $\mathbf{n}$ is

$$\sigma_N = \mathbf{t}^{(\mathbf{n})} \cdot \mathbf{n} = (\boldsymbol{\sigma}\mathbf{n}) \cdot \mathbf{n} \tag{3.4.7}$$

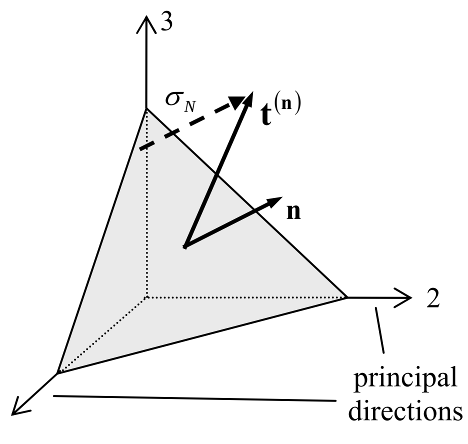


Figure 3.4.2: normal stress acting on a plane defined by the unit normal n

With respect to the principal stresses, using 3.4.6,

$$\mathbf{t}^{(\mathbf{n})} = \boldsymbol{\sigma}\,\mathbf{n} = \sigma_1 n_1 \mathbf{e}_1 + \sigma_2 n_2 \mathbf{e}_2 + \sigma_3 n_3 \mathbf{e}_3 \tag{3.4.8}$$

and the normal stress is

$$\sigma_N = \sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2 \tag{3.4.9}$$

Since $n_1^2 + n_2^2 + n_3^2 = 1$ and, without loss of generality, taking $\sigma_1 \ge \sigma_2 \ge \sigma_3$, one has

$$\sigma_1 = \sigma_1\left(n_1^2 + n_2^2 + n_3^2\right) \ge \sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2 = \sigma_N \tag{3.4.10}$$

Similarly,

$$\sigma_N = \sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2 \ge \sigma_3\left(n_1^2 + n_2^2 + n_3^2\right) \ge \sigma_3 \tag{3.4.11}$$

Thus the maximum normal stress acting at a point is the maximum principal stress and the minimum normal stress acting at a point is the minimum principal stress.

Shear Stresses

Next, it will be shown that the maximum shearing stresses at a point act on planes oriented at 45° to the principal planes and that they have magnitude equal to half the difference between the principal stresses.

From 3.3.39, 3.4.8 and 3.4.9, the shear stress on the plane is

$$\sigma_S^2 = \left(\sigma_1^2 n_1^2 + \sigma_2^2 n_2^2 + \sigma_3^2 n_3^2\right) - \left(\sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2\right)^2 \tag{3.4.12}$$

Using the condition $n_1^2 + n_2^2 + n_3^2 = 1$ to eliminate n_3 leads to

$$\sigma_S^2 = \left(\sigma_1^2 - \sigma_3^2\right)n_1^2 + \left(\sigma_2^2 - \sigma_3^2\right)n_2^2 + \sigma_3^2 - \left[\left(\sigma_1 - \sigma_3\right)n_1^2 + \left(\sigma_2 - \sigma_3\right)n_2^2 + \sigma_3\right]^2 \tag{3.4.13}$$

The stationary points are now obtained by equating the partial derivatives with respect to the two variables n_1 and n_2 to zero:

$$\begin{aligned}
\frac{\partial\left(\sigma_S^2\right)}{\partial n_1} &= n_1\left(\sigma_1 - \sigma_3\right)\left\{\sigma_1 - \sigma_3 - 2\left[\left(\sigma_1 - \sigma_3\right)n_1^2 + \left(\sigma_2 - \sigma_3\right)n_2^2\right]\right\} = 0 \\
\frac{\partial\left(\sigma_S^2\right)}{\partial n_2} &= n_2\left(\sigma_2 - \sigma_3\right)\left\{\sigma_2 - \sigma_3 - 2\left[\left(\sigma_1 - \sigma_3\right)n_1^2 + \left(\sigma_2 - \sigma_3\right)n_2^2\right]\right\} = 0
\end{aligned} \tag{3.4.14}$$

One sees immediately that $n_1 = n_2 = 0$ (so that $n_3 = \pm 1$) is a solution; this is the principal direction $\mathbf{e}_3$ and the shear stress is by definition zero on the plane with this normal. In this calculation, the component n_3 was eliminated and σ_S^2 was treated as a function of the variables (n_1, n_2). Similarly, n_1 can be eliminated with (n_2, n_3) treated as the variables, leading to the solution $\mathbf{n} = \mathbf{e}_1$, and n_2 can be eliminated with (n_1, n_3) treated as the variables, leading to the solution $\mathbf{n} = \mathbf{e}_2$. Thus these solutions lead to the minimum shear stress value $\sigma_S^2 = 0$.

A second solution to Eqn. 3.4.14 can be seen to be $n_1 = 0, n_2 = \pm 1/\sqrt{2}$ (so that $n_3 = \pm 1/\sqrt{2}$) with corresponding shear stress values $\sigma_S^2 = \frac{1}{4}(\sigma_2 - \sigma_3)^2$. Two other solutions can be obtained as described earlier, by eliminating n_1 and by eliminating n_2. The full solution is listed below, and these are evidently the maximum (absolute value of the) shear stresses acting at a point:

$$
\begin{aligned}
\mathbf{n} &= \left(0, \pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right), \quad \sigma_S = \frac{1}{2}|\sigma_2 - \sigma_3| \\
\mathbf{n} &= \left(\pm\frac{1}{\sqrt{2}}, 0, \pm\frac{1}{\sqrt{2}}\right), \quad \sigma_S = \frac{1}{2}|\sigma_3 - \sigma_1| \\
\mathbf{n} &= \left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, 0\right), \quad \sigma_S = \frac{1}{2}|\sigma_1 - \sigma_2|
\end{aligned}
\tag{3.4.15}
$$

Taking $\sigma_1 \ge \sigma_2 \ge \sigma_3$, the maximum shear stress at a point is

$$\tau_{\max} = \frac{1}{2}(\sigma_1 - \sigma_3) \tag{3.4.16}$$

and acts on a plane with normal oriented at 45° to the 1 and 3 principal directions. This is illustrated in Fig. 3.4.3.

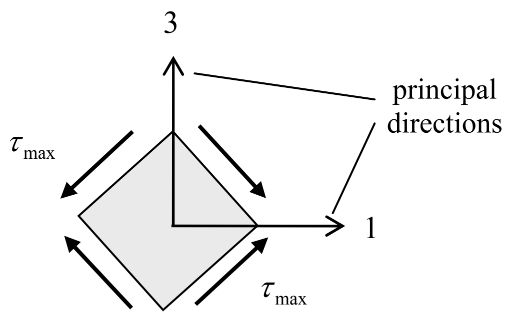


Figure 3.4.3: maximum shear stress at apoint

Example (maximum shear stress)

Consider the stress state

$$[\sigma_{ij}] = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -6 & -12 \\ 0 & -12 & 1 \end{bmatrix}$$

This is the same tensor considered in the example of §1.11.1. Using the results of that example, the principal stresses are $\sigma_1 = 10, \sigma_2 = 5, \sigma_3 = -15$ and so the maximum shear stress at that point is

$$\tau_{\max} = \frac{1}{2}(\sigma_1 - \sigma_3) = \frac{25}{2}$$

The planes and direction upon which they act are shown in Fig. 3.4.4.

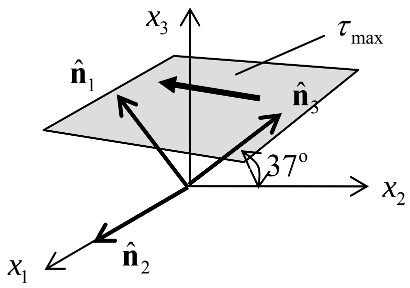


Figure 3.4.4: maximum shear stress

■

3.5 Stress Measures for Large Deformations

Thus far, the surface forces acting within a material have been described in terms of the Cauchy stress tensor $\boldsymbol{\sigma}$. The Cauchy stress is also called the **true stress**, to distinguish it from other stress tensors, some of which will be discussed below. It is called the *true* stress because it is a true measure of the force per unit area in the current, deformed, configuration. When the deformations are small, there is no distinction to be made between this deformed configuration and some reference, or undeformed, configuration, and the Cauchy stress is the sensible way of describing the action of surface forces. When the deformations are large, however, one needs to refer to some reference configuration. In this case, there are a number of different possible ways of defining the action of surface forces; some of these stress measures often do not have as clear a physical meaning as the Cauchy stress, but are useful nonetheless.

3.5.1 The First Piola – Kirchhoff Stress Tensor

Consider two configurations of a material, the reference and current configurations. Consider now a vector element of surface in the reference configuration, $\mathbf{N}dS$, where dS is the area of the element and $\mathbf{N}$ is the unit normal. After deformation, the material particles making up this area element now occupy the element defined by $\mathbf{n}ds$, where ds is the area and $\mathbf{n}$ is the normal in the current configuration. Suppose that a force $d\mathbf{f}$ acts on the surface element (in the current configuration). Then by definition of the Cauchy stress

$$d\mathbf{f} = \boldsymbol{\sigma}\,\mathbf{n}\,ds \tag{3.5.1}$$

The **first Piola-Kirchhoff stress** tensor $\mathbf{P}$ (which will be called the **PK1 stress** for brevity) is defined by

$$d\mathbf{f} = \mathbf{P}\,\mathbf{N}\,dS \tag{3.5.2}$$

The PK1 stress relates the force acting in the *current* configuration to the surface element in the *reference* configuration. Since it relates to both configurations, it is a two-point tensor.

The (Cauchy) traction vector was defined as

$$\mathbf{t} = \frac{d\mathbf{f}}{ds}, \qquad \mathbf{t} = \boldsymbol{\sigma}\,\mathbf{n} \tag{3.5.3}$$

Similarly, one can introduce a **PK1 traction vector T** such that

$$\mathbf{T} = \frac{d\mathbf{f}}{dS}, \qquad \mathbf{T} = \mathbf{P}\,\mathbf{N} \tag{3.5.4}$$

Whereas the Cauchy traction is the actual physical force per area on the element in the current configuration, the PK1 traction is a fictitious quantity – the force acting on an element in the current configuration divided by the area of the corresponding element in the reference configuration. Note that, since $d\mathbf{f} = \mathbf{t}ds = \mathbf{T}dS$, it follows that $\mathbf{T}$ and $\mathbf{t}$ act in the same direction (but have different magnitudes), Fig. 3.5.1.

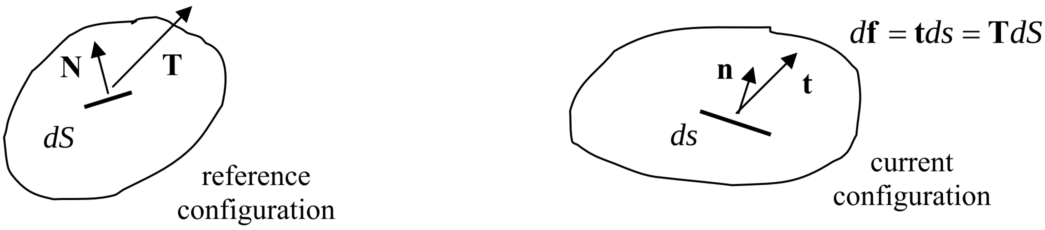


Figure 3.5.1: Traction vectors

Uniaxial Tension

Consider a uniaxial tensile test whereby a specimen is stretched uniformly by a constant force $\mathbf{f}$, Fig. 3.5.2. The initial cross-sectional area of the specimen is A_0 and the cross-sectional area of the specimen at time t is $A(t)$. The Cauchy (true) stress is

$$\boldsymbol{\sigma}(t) = \frac{\mathbf{f}}{A(t)} \tag{3.5.5}$$

and the PK1 stress is

$$\mathbf{P} = \frac{\mathbf{f}}{A_0} \tag{3.5.6}$$

This stress measure, force over area of the undeformed specimen, as used in the uniaxial tensile test, is also called the **engineering stress**.

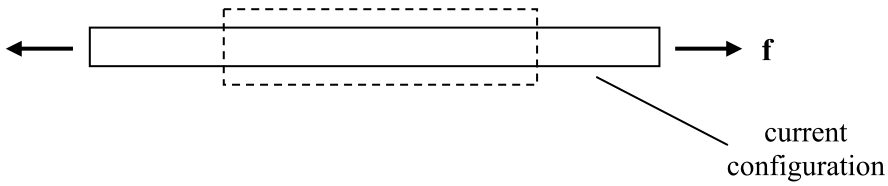


Figure 3.5.2: Uniaxial tension of a bar

The Nominal Stress

The PK1 stress tensor is also called the **nominal stress tensor**. Note that many authors use a different definition for the nominal stress, namely $\mathbf{T} = \mathbf{N}\mathbf{P}$, and then define the PK1 stress to be the transpose of this $\mathbf{P}$. Thus all authors use the same definition for the PK1 stress, but a slightly different definition for the nominal stress.

Relation between the Cauchy and PK1 Stresses

From the above definitions,

$$\boldsymbol{\sigma}\mathbf{n}ds = \mathbf{P}\mathbf{N}dS \tag{3.5.7}$$

Using Nanson's formula, 2.2.59, $\mathbf{n}ds = J\mathbf{F}^{-\mathrm{T}}\mathbf{N}dS$,

$$\boxed{\begin{aligned}\mathbf{P} &= J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}} \\ \boldsymbol{\sigma} &= J^{-1}\mathbf{P}\mathbf{F}^{\mathrm{T}}\end{aligned}} \qquad \textbf{PK1 stress} \tag{3.5.8}$$

The Cauchy stress is symmetric, but the deformation gradient is not. Hence the PK1 stress tensor is *not symmetric*, and this restricts its use as an alternative stress measure to the Cauchy stress measure. In fact, this lack of symmetry and lack of a clear physical meaning makes it uncommon for the PK1 stress to be used in the modeling of materials. It is, however, useful in the description of the momentum balance laws in the material description, where **P** plays an analogous role to that played by the Cauchy stress $\boldsymbol{\sigma}$ in the equations of motion (see later).

3.5.2 The Second Piola – Kirchhoff Stress Tensor

The **second Piola – Kirchhoff stress** tensor, or the **PK2 stress**, **S**, is defined by

$$\boxed{\mathbf{S} = J\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}} \qquad \textbf{PK2 stress} \tag{3.5.9}$$

Even though the PK2 does not admit a physical interpretation (except in the simplest of cases, but see the interpretation below), there are three good reasons for using it as a measure of the forces acting in a material. First, one can see that

$$\left(\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right)^{\mathrm{T}} = \left(\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right)^{\mathrm{T}}\left(\mathbf{F}^{-1}\right)^{\mathrm{T}} = \mathbf{F}^{-1}\boldsymbol{\sigma}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}}$$

and since the Cauchy stress is symmetric, so is the PK2 stress:

$$\mathbf{S} = \mathbf{S}^{\mathrm{T}} \tag{3.5.10}$$

A second reason for using the PK2 stress is that, together with the Euler-Lagrange strain **E**, it gives the power of a deforming material (see later). Third, it is parameterized by material coordinates only, that is, it is a material tensor field, in the same way as the Cauchy stress is a spatial tensor field.

Note that the PK1 and PK2 stresses are related through

$$\mathbf{P} = \mathbf{F}\mathbf{S}, \qquad \mathbf{S} = \mathbf{F}^{-1}\mathbf{P} \tag{3.5.11}$$

The PK2 stress can be interpreted as follows: take the force vector in the current configuration $d\mathbf{f}$ and locate a corresponding vector in the undeformed configuration according to $d\bar{\mathbf{f}} = \mathbf{F}^{-1}d\mathbf{f}$. The PK2 stress tensor is this fictitious force divided by the corresponding area element in the reference configuration: $d\bar{\mathbf{f}} = \mathbf{SN}dS$, and 3.5.9 follows from 3.5.2, 3.5.8:

$$d\mathbf{f} = \mathbf{P}\,\mathbf{N}dS = J\boldsymbol{\sigma}\,\mathbf{F}^{-\mathrm{T}}\,\mathbf{N}dS$$

3.5.3 Alternative Stress Tensors

Some other useful stress measures are described here.

The Kirchhoff Stress

The **Kirchhoff stress tensor** $\boldsymbol{\tau}$ is defined as

$$\boxed{\boldsymbol{\tau} = J\boldsymbol{\sigma}} \qquad \textbf{Kirchhoff Stress} \qquad (3.5.12)$$

It is a spatial tensor field parameterized by spatial coordinates. One reason for its use is that, in many equations, the Cauchy stress appears together with the Jacobian and the use of $\boldsymbol{\tau}$ simplifies formulae.

Note that the Kirchhoff stress is the push forward of the PK2 stress; from 2.12.9b, 2.12.11b,

$$\begin{aligned} \boldsymbol{\tau} &= \chi_*(\mathbf{S})^\# = \mathbf{FSF}^{\mathrm{T}} \\ \mathbf{S} &= \chi_*^{-1}(\boldsymbol{\tau})^\# = \mathbf{F}^{-1}\boldsymbol{\tau}\mathbf{F}^{-\mathrm{T}} \end{aligned} \qquad (3.5.13)$$

The Corotational Cauchy Stress

The **corotational stress** $\hat{\boldsymbol{\sigma}}$ is defined as

$$\boxed{\hat{\boldsymbol{\sigma}} = \mathbf{R}^{\mathrm{T}}\boldsymbol{\sigma}\mathbf{R}} \qquad \textbf{Corotational Stress} \qquad (3.5.14)$$

where **R** is the orthogonal rotation tensor. Whereas the Cauchy stress is related to the PK2 stress through $\boldsymbol{\sigma} = J^{-1}\mathbf{FSF}^{\mathrm{T}}$, the corotational stress is related to the PK2 stress through (with **F** replaced by the right (symmetric) stretch tensor **U**):

$$\hat{\boldsymbol{\sigma}} = J^{-1}\mathbf{USU}^{\mathrm{T}} = J^{-1}\mathbf{U}\left(J\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right)\mathbf{U} = \left(\mathbf{UF}^{-1}\right)\boldsymbol{\sigma}\left(\mathbf{F}^{-\mathrm{T}}\mathbf{U}\right) = \mathbf{R}^{\mathrm{T}}\boldsymbol{\sigma}\mathbf{R} \qquad (3.5.15)$$

The corotational stress is defined on the intermediate configuration of Fig. 2.10.8. It can be regarded as the push forward of the PK2 stress from the reference configuration through the stretch **U**, scaled by J^{-1} (Eqn. 2.12.28b):

$$\hat{\boldsymbol{\sigma}} = J^{-1}\chi_*(\mathbf{S})^\#{}_{\mathrm{U(G)}} = J^{-1}S^{ij}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}_j = J^{-1}S^{ij}\left(\mathbf{UG}_i \otimes \mathbf{UG}_j\right) = J^{-1}\mathbf{USU}^{\mathrm{T}} = J^{-1}\mathbf{USU} \qquad (3.5.16)$$

or as the pull-back of the Cauchy stress with respect to **R** (Eqn. 2.12.27f):

$$\hat{\boldsymbol{\sigma}} = \chi_*^{-1}(\boldsymbol{\sigma})^{\#}{}_{\mathbf{R}(\mathbf{g})} = \sigma^{ij}\hat{\mathbf{g}}_i \otimes \hat{\mathbf{g}}_j = \mathbf{R}^{\mathrm{T}}\boldsymbol{\sigma}\mathbf{R} \qquad (3.5.17)$$

The Biot Stress

The **Biot** (or **Jaumann**) **stress tensor** $\mathbf{T}_{\mathrm{B}}$ is defined as

$$\boxed{\mathbf{T}_{\mathrm{B}} = \mathbf{R}^{\mathrm{T}}\mathbf{P} = \mathbf{U}\mathbf{S}} \qquad \textbf{Biot Stress} \qquad (3.5.18)$$

From 3.5.11, it is similar to the PK1 stress, only with **F** replaced by **U**.

Example

Consider a **pre-stressed** thin plate with $\sigma_{11} = \sigma_1^0$, $\sigma_{22} = \sigma_2^0$, that is, it has a non-zero stress although no forces are acting[1], Fig. 3.5.3. In this initial state, $\mathbf{F} = \mathbf{I}$ and, considering a two-dimensional state of stress,

$$\boldsymbol{\sigma} = \mathbf{P} = \mathbf{S} = \hat{\boldsymbol{\sigma}} = \boldsymbol{\tau} = \mathbf{T}_{\mathrm{B}} = \begin{bmatrix} \sigma_1^0 & 0 \\ 0 & \sigma_2^0 \end{bmatrix}$$

The material is now rotated *as a rigid body* 45° counterclockwise – the stress-state is "frozen" within the material and rotates with it. Then

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

The stress components with respect to the rotated x_i^* axes shown in Fig. 3.5.3b are $\sigma_{11}^* = \sigma_1^0$, etc.; the components with respect to the spatial axes x_i can be found from the stress transformation rule $[\boldsymbol{\sigma}] = [\mathbf{Q}^{\mathrm{T}}][\boldsymbol{\sigma}^*][\mathbf{Q}] = [\mathbf{R}][\boldsymbol{\sigma}^*][\mathbf{R}^{\mathrm{T}}]$, and so

$$\boldsymbol{\sigma} = \begin{bmatrix} \frac{1}{2}(\sigma_1^0 + \sigma_2^0) & \frac{1}{2}(\sigma_1^0 - \sigma_2^0) \\ \frac{1}{2}(\sigma_1^0 - \sigma_2^0) & \frac{1}{2}(\sigma_1^0 + \sigma_2^0) \end{bmatrix}$$

Note that the Cauchy stress changes with this rigid body rotation. Further, with $J = 1$,

$$\boldsymbol{\tau} = \boldsymbol{\sigma}, \quad \mathbf{P} = \begin{bmatrix} \sigma_1^0/\sqrt{2} & -\sigma_2^0/\sqrt{2} \\ \sigma_1^0/\sqrt{2} & \sigma_2^0/\sqrt{2} \end{bmatrix}, \quad \mathbf{S} = \hat{\boldsymbol{\sigma}} = \mathbf{T}_{\mathrm{B}} = \begin{bmatrix} \sigma_1^0 & 0 \\ 0 & \sigma_2^0 \end{bmatrix}$$

Note that the PK1 stress is not symmetric. Now attach axes x^* to the material and rotate these axes with the specimen as it rotates, as in Fig. 3.5.3b. The components with respect

[1] for example a piece of metal can be deformed; when the *load is removed* it is often pre-stressed – there is a non-zero state of stress in the material

to these rotated axes give the corotational stress; the corotational stress is the stress in a body, taking out the stress changes caused by rigid body rotations – one says that the corotational stress (and PK2 stress) "rotate" with the body.

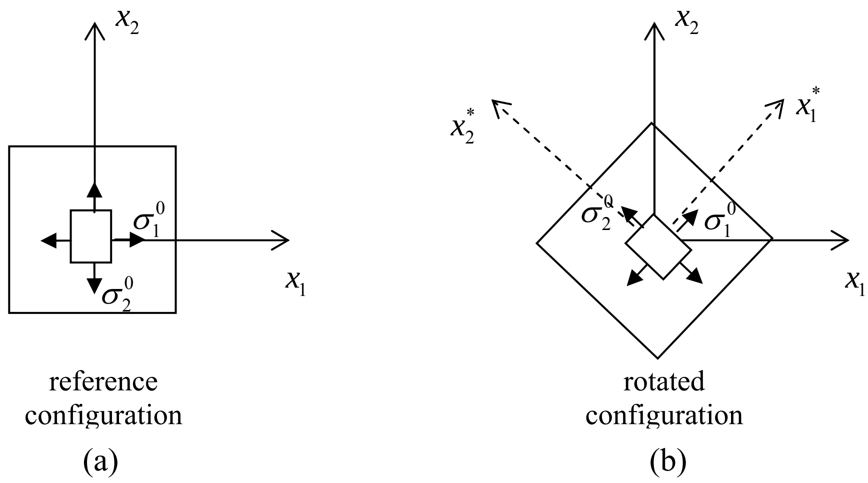


Figure 3.5.3: Pre-stressed material; (a) original position, (b) rotated configuration

■

3.5.4 Small deformations

From §2.7, when the deformations are small, neglecting terms involving products of displacement gradients,

$$\mathbf{F} = \mathbf{I} + \mathrm{grad}\mathbf{u} + \mathrm{O}(\mathrm{grad}\mathbf{u})^2 = \mathbf{I} + \mathrm{O}(\mathrm{grad}\mathbf{u}) \tag{3.5.19}$$

Here, $\mathrm{O}(\mathrm{grad}\mathbf{u})$ means terms of the order of displacement gradients (and higher) have been neglected and $\mathrm{O}(\mathrm{grad}\mathbf{u})^2$ means terms of the order of products of displacement gradients (and higher) have been neglected. Also,

$$\begin{aligned} J &= \det\mathbf{F} \\ &= \det\left(\mathbf{I} + \mathrm{grad}\mathbf{u} + \mathrm{O}(\mathrm{grad}\mathbf{u})^2\right) = 1 + \mathrm{div}\mathbf{u} + \mathrm{O}(\mathrm{grad}\mathbf{u})^2 = 1 + \mathrm{O}(\mathrm{grad}\mathbf{u}) \end{aligned} \tag{3.5.20}$$

From 3.5.8 and 3.5.9, using 3.5.19-20, one has

$$\begin{aligned} J\boldsymbol{\sigma} &= \mathbf{P}\mathbf{F}^{\mathrm{T}} \quad \rightarrow \quad \boldsymbol{\sigma} + \mathrm{O}(\mathrm{grad}\mathbf{u}) = \mathbf{P} + \mathrm{O}(\mathrm{grad}\mathbf{u}) \\ J\boldsymbol{\sigma} &= \mathbf{F}\mathbf{S}\mathbf{F}^{\mathrm{T}} \quad \rightarrow \quad \boldsymbol{\sigma} + \mathrm{O}(\mathrm{grad}\mathbf{u}) = \mathbf{S} + \mathrm{O}(\mathrm{grad}\mathbf{u}) \end{aligned} \tag{3.5.21}$$

In the linear theory then, with $\mathrm{O}(\mathrm{grad}\mathbf{u}) \rightarrow 0$, the stress measures encountered in this section are all equivalent.

3.5.5 Objective Stress Tensors

In order to ascertain the objectivity of the stress tensors, first note that, *by definition*, force is an objective vector, and therefore so also is the traction vector. Similarly for the normal vector. The normal and traction vectors transform under an observer transformation according to 2.8.10, $\mathbf{n}^* = \mathbf{Qn}$ and $\mathbf{t}^* = \mathbf{Qt}$. Then

$$\mathbf{t} = \boldsymbol{\sigma}\mathbf{n} \quad \rightarrow \quad \mathbf{Q}^{\mathrm{T}}\mathbf{t}^* = \boldsymbol{\sigma}\mathbf{Q}^{\mathrm{T}}\mathbf{n}^* \quad \rightarrow \quad \mathbf{t}^* = \left(\mathbf{Q}\boldsymbol{\sigma}\mathbf{Q}^{\mathrm{T}}\right)\mathbf{n}^* \tag{3.5.22}$$

and so $\boldsymbol{\sigma}^* = \mathbf{Q}\boldsymbol{\sigma}\mathbf{Q}^{\mathrm{T}}$; according to 2.8.12, the Cauchy stress is objective. The PK2 stress $\mathbf{S}$ is objective, since it is a material tensor unaffected by an observer transformation. For the PK1 stress, using 2.8.23,

$$\mathbf{P}^* = J^*\boldsymbol{\sigma}^*\left(\mathbf{F}^*\right)^{-\mathrm{T}} = J\mathbf{Q}\boldsymbol{\sigma}\mathbf{Q}^{\mathrm{T}}\left(\mathbf{QF}\right)^{-\mathrm{T}} = \mathbf{Q}\left(J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right) \tag{3.5.23}$$

and so, according to 2.8.16, $\mathbf{P}$ is objective (transforming like a vector, being a two-point tensor).

3.5.6 Objective Stress Rates

One needs to incorporate stress rates in models of materials where the response depends on the rate of stressing, for example with viscoelastic materials. As discussed in §2.8.5, the rates of objective tensors are not necessarily objective. As discussed in §2.12.3, the Lie derivative of a spatial second order tensor is objective. For the Cauchy stress, there are a number of different objective rates one can use, based on the Lie derivative (see Eqns. 2.8.35-36, 2.12.41, 2.12.44):

$$\begin{array}{lll} \textbf{Cotter-Rivlin stress rate} & \dot{\boldsymbol{\sigma}} + \mathbf{l}^{\mathrm{T}}\boldsymbol{\sigma} + \boldsymbol{\sigma}\mathbf{l} & = \mathrm{L}_{\mathrm{v}}^{\mathrm{b}}\boldsymbol{\sigma} \\ \textbf{Jaumann stress rate} & \dot{\boldsymbol{\sigma}} - \mathbf{w}\boldsymbol{\sigma} + \boldsymbol{\sigma}\mathbf{w} & = \frac{1}{2}\left(\mathrm{L}_{\mathrm{v}}^{\mathrm{b}}\boldsymbol{\sigma} + \mathrm{L}_{\mathrm{v}}^{\#}\boldsymbol{\sigma}\right) \\ \textbf{Oldroyd stress rate}^{2} & \dot{\boldsymbol{\sigma}} - \mathbf{l}\boldsymbol{\sigma} - \boldsymbol{\sigma}\mathbf{l}^{\mathrm{T}} & = \mathrm{L}_{\mathrm{v}}^{\#}\boldsymbol{\sigma} \end{array} \tag{3.5.24}$$

Stress rates of other spatial stress tensors can be defined in the same way, for example the Oldroyd rate of the Kirchhoff stress tensor is $\dot{\boldsymbol{\tau}} - \mathbf{l}\boldsymbol{\tau} - \boldsymbol{\tau}\mathbf{l}^{\mathrm{T}}$.

The material derivative of the material PK2 stress tensor, $\dot{\mathbf{S}}$, is objective. The push forward of $\dot{\mathbf{S}}$ is, from 2.12.9b,

$$\chi_*\left(\dot{\mathbf{S}}\right)^{\#} = \mathbf{F}\dot{\mathbf{S}}\mathbf{F}^{\mathrm{T}} \tag{3.5.25}$$

---

[2] this is sometimes called the contravariant Oldroyd stress rate, to distinguish it from the Cotter-Rivlin rate, which is also sometimes called the covariant Oldroyd stress rate

This push forward, scaled by the inverse of the Jacobian, $J^{-1}\mathbf{F}\dot{\mathbf{S}}\mathbf{F}^{\mathrm{T}}$ is called the **Truesdell stress rate**. This can be expressed in terms of the Cauchy stress by using 3.5.9, and then 2.5.20, 2.5.5:

$$\begin{aligned} J^{-1}\mathbf{F}\frac{d}{dt}\left(J\mathbf{F}^{-1}\boldsymbol{\sigma}\,\mathbf{F}^{-\mathrm{T}}\right)\mathbf{F}^{\mathrm{T}} &= J^{-1}\mathbf{F}\left(\dot{J}\mathbf{F}^{-1}\boldsymbol{\sigma}\,\mathbf{F}^{-\mathrm{T}} + J\,\dot{\overline{\mathbf{F}^{-1}}}\,\boldsymbol{\sigma}\,\mathbf{F}^{-\mathrm{T}} + J\mathbf{F}^{-1}\dot{\boldsymbol{\sigma}}\,\mathbf{F}^{-\mathrm{T}} + J\mathbf{F}^{-1}\boldsymbol{\sigma}\,\dot{\overline{\mathbf{F}^{-\mathrm{T}}}}\right)\mathbf{F}^{\mathrm{T}} \\ &= \dot{\boldsymbol{\sigma}} - \mathbf{l}\boldsymbol{\sigma} - \boldsymbol{\sigma}\mathbf{l}^{\mathrm{T}} + \mathrm{tr}(\mathbf{d})\boldsymbol{\sigma} \end{aligned} \tag{3.5.26}$$

Thus far, objective rates have been constructed by pulling back, taking derivatives and pushing forward. One can construct objective rates also by pulling back and pushing forward with the rotation tensor **R** only, since it is the rotation which causes the stress rates to be non-objective. For example, $\mathrm{L}_{\mathrm{v}}^{\#}\boldsymbol{\sigma}$, setting $\mathbf{F}=\mathbf{R}$, is, from 3.5.17 and 2.12.27b,

$$\begin{aligned} \chi_*\left(\frac{d}{dt}\left[\chi_*^{-1}(\boldsymbol{\sigma})^{\#}_{\mathbf{R}(\mathbf{g})}\right]\right)_{\mathbf{R}(\hat{\mathbf{g}})} &= \chi_*\left(\frac{d}{dt}[\hat{\boldsymbol{\sigma}}]\right)^{\#}_{\mathbf{R}(\hat{\mathbf{g}})} \\ &= \mathbf{R}\left(\dot{\mathbf{R}}^{\mathrm{T}}\boldsymbol{\sigma}\,\mathbf{R} + \mathbf{R}^{\mathrm{T}}\dot{\boldsymbol{\sigma}}\,\mathbf{R} + \mathbf{R}^{\mathrm{T}}\boldsymbol{\sigma}\,\dot{\mathbf{R}}\right)\mathbf{R}^{\mathrm{T}} \\ &= \dot{\boldsymbol{\sigma}} + \boldsymbol{\sigma}\boldsymbol{\Omega}_R - \boldsymbol{\Omega}_R\boldsymbol{\sigma} \end{aligned} \tag{3.5.27}$$

where $\boldsymbol{\Omega}_R = \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}}$ is the skew-symmetric angular velocity tensor 2.6.3. The stress rate 3.5.27 is called the **Green-Naghdi stress rate**. From the above, the Green-Naghdi rate is the push forward of the time derivative of the corotational stress.

Example

Consider again the example discussed at the end of §3.5.3, only let the plate rotate at constant angular velocity ω, so

$$\mathbf{F}=\mathbf{R}=\begin{bmatrix}\cos(\omega t) & -\sin(\omega t)\\ \sin(\omega t) & \cos(\omega t)\end{bmatrix},\quad \dot{\mathbf{F}}=\dot{\mathbf{R}}=\omega\begin{bmatrix}-\sin(\omega t) & -\cos(\omega t)\\ \cos(\omega t) & -\sin(\omega t)\end{bmatrix}$$

Again, using the stress transformation rule $[\boldsymbol{\sigma}]=[\mathbf{Q}^{\mathrm{T}}][\boldsymbol{\sigma}^*][\mathbf{Q}]=[\mathbf{R}][\boldsymbol{\sigma}^*][\mathbf{R}^{\mathrm{T}}]$,

$$\boldsymbol{\sigma}=\begin{bmatrix}\sigma_1^0\cos^2(\omega t)+\sigma_2^0\sin^2(\omega t) & \cos(\omega t)\sin(\omega t)\left(\sigma_1^0-\sigma_2^0\right)\\ \cos(\omega t)\sin(\omega t)\left(\sigma_1^0-\sigma_2^0\right) & \sigma_1^0\sin^2(\omega t)+\sigma_2^0\cos^2(\omega t)\end{bmatrix}$$

and, with $J=1$,

$$\boldsymbol{\tau}=\boldsymbol{\sigma},\quad \mathbf{P}=\begin{bmatrix}\cos(\omega t)\sigma_1^0 & -\sin(\omega t)\sigma_2^0\\ \sin(\omega t)\sigma_1^0 & \cos(\omega t)\sigma_2^0\end{bmatrix},\quad \mathbf{S}=\hat{\boldsymbol{\sigma}}=\mathbf{T}_{\mathrm{B}}=\begin{bmatrix}\sigma_1^0 & 0\\ 0 & \sigma_2^0\end{bmatrix}$$

Also,

$$\mathbf{l} = \mathbf{w} = \dot{\mathbf{F}}\mathbf{F}^{-1} = \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}} = \mathbf{\Omega}_{\mathbf{R}} = \omega\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Then

$$\dot{\mathbf{P}} = \omega\begin{bmatrix} -\sin(\omega t)\sigma_1^0 & -\cos(\omega t)\sigma_2^0 \\ \cos(\omega t)\sigma_1^0 & -\sin(\omega t)\sigma_2^0 \end{bmatrix}, \quad \dot{\mathbf{S}} = \dot{\hat{\boldsymbol{\sigma}}} = \dot{\mathbf{T}}_{\mathrm{B}} = \mathbf{0}$$

and

$$\dot{\boldsymbol{\sigma}} = \omega\begin{bmatrix} -2\sin(\omega t)\cos(\omega t)\left(\sigma_1^0 - \sigma_2^0\right) & \left(\cos^2(\omega t) - \sin^2(\omega t)\right)\left(\sigma_1^0 - \sigma_2^0\right) \\ \left(\cos^2(\omega t) - \sin^2(\omega t)\right)\left(\sigma_1^0 - \sigma_2^0\right) & +2\sin(\omega t)\cos(\omega t)\left(\sigma_1^0 - \sigma_2^0\right) \end{bmatrix}$$

For a rigid body rotation, it can be seen that the definitions of the Cotter-Rivlin, Jaumann, Oldroyd, Truesdell and Green-Naghdi rates are equivalent, and they are all zero:

$$\dot{\boldsymbol{\sigma}} - \mathbf{w}\boldsymbol{\sigma} + \boldsymbol{\sigma}\mathbf{w} = \mathbf{0}$$

This is as expected since objective stress rates for two configurations which differ by a rigid body rotation will, by definition, be equal (the stress components will not change); they are zero in the reference configuration and so will be zero in the rotated configuration.

■

3.5.7 Problems

1. Consider the case of uniaxial stress, where a material with initial dimensions length l_0, breadth w_0 and height h_0 deforms into a component with dimensions length l, breadth w and height h. The only non-zero Cauchy stress component is σ_{11}, acting in the direction of the length of the component.
   (a) write down the motion equations in the material description, $\mathbf{x} = \chi(\mathbf{X})$
   (b) calculate the deformation gradient $\mathbf{F}$ and confirm that $J = \det\mathbf{F}$ is the ratio of the volume in the current configuration to that in the initial configuration
   (c) Calculate the PK1 stress. How is it related to the Cauchy stress for this uniaxial stress-state?
   (d) calculate the PK2 stress
2. A material undergoes the deformation
   $$x_1 = 3X_1 t, \quad x_2 = X_1 t + X_2, \quad x_3 = X_3$$
   The Cauchy stress at a point in the material is
   $$[\boldsymbol{\sigma}] = \begin{bmatrix} t & -2t & 0 \\ -2t & t & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
   (a) Calculate the PK1 and PK2 stresses at the point (check that PK2 is symmetric)

(b) Calculate the expressions $\mathbf{P} : \dot{\mathbf{F}}$, $J\boldsymbol{\sigma} : \mathbf{d}$, $\mathbf{S} : \dot{\mathbf{E}}$ (for $\dot{\mathbf{E}}$, use the expression 2.5.18b, $\dot{\mathbf{E}} = \mathbf{F}^{\mathrm{T}}\mathbf{dF}$). In these expressions, $\mathbf{d}$ is the rate of deformation tensor. (You should get the same result for all three cases, since they all give the rate of internal work done by the stresses during the deformation, per unit reference volume – see later)

3. Show that the Oldroyd rate of the Kirchhoff stress, $\dot{\boldsymbol{\tau}} - \mathbf{l}\boldsymbol{\tau} - \boldsymbol{\tau}\mathbf{l}^{\mathrm{T}}$, is equal to the Jacobian times the Truesdell stress rate of the Cauchy stress, 3.5.26.

3.6 The Equations of Motion and Symmetry of Stress

In Part II, §1.1, the Equations of Motion were derived using Newton's Law applied to a differential material element. Here, they are derived using the principle of linear momentum.

3.6.1 The Equations of Motion (Spatial Form)

Application of Cauchy's law $\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}$ and the divergence theorem 1.14.21 to 3.2.7 leads directly to the global form of the equations of motion

$$\int_v [\operatorname{div}\boldsymbol{\sigma} + \mathbf{b}]dv = \int_v \rho\,\dot{\mathbf{v}}\,dv, \qquad \int_v \left[\frac{\partial \sigma_{ij}}{\partial x_j} + b_i\right]dv = \int_v \rho\,\dot{v}_i\,dv \tag{3.6.1}$$

The corresponding local form is then

$$\boxed{\operatorname{div}\boldsymbol{\sigma} + \mathbf{b} = \rho\frac{d\mathbf{v}}{dt}, \qquad \frac{\partial \sigma_{ij}}{\partial x_j} + b_i = \rho\frac{dv_i}{dt}} \quad \textbf{Equations of Motion} \tag{3.6.2}$$

The term on the right is called the inertial, or kinetic, term, representing the change in momentum. The material time derivative of the spatial velocity field is

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + (\operatorname{grad}\mathbf{v})\mathbf{v} \quad \text{so} \quad \frac{dv_1}{dt} = \frac{\partial v_1}{\partial t} + \left(\frac{\partial v_1}{\partial x_1}v_1 + \frac{\partial v_1}{\partial x_2}v_2 + \frac{\partial v_1}{\partial x_3}v_3\right), \text{ etc.}$$

and it can be seen that the equations of motion are non-linear in the velocities.

Equations of Equilibrium

When the acceleration is zero, the equations reduce to the equations of equilibrium,

$$\boxed{\operatorname{div}\boldsymbol{\sigma} + \mathbf{b} = \mathbf{0}} \quad \textbf{Equations of Equilibrium} \tag{3.6.3}$$

Flows

A **flow** is a set of quantities associated with the system of forces **t** and **b**, for example the quantities $\mathbf{v}, \boldsymbol{\sigma}, \rho$. A flow is **steady** if the associated spatial quantities are independent of time. A **potential flow** is one for which the velocity field can be written as the gradient of a scalar function, $\mathbf{v} = \operatorname{grad}\phi$. An **irrotational flow** is one for which $\operatorname{curl}\mathbf{v} = \mathbf{0}$.

3.6.2 The Equations of Motion (Material Form)

In the spatial form, the linear momentum of a mass element is $\rho \mathbf{v} dv$. In the material form it is $\rho_0 \mathbf{V} dV$. Here, **V** is the same velocity as **v**, only it is now expressed in terms of the material coordinates **X**, and $\rho dv = \rho_0 dV$. The linear momentum of a collection of material particles occupying the volume *v* in the current configuration can thus be expressed in terms of an integral over the corresponding volume *V* in the reference configuration:

$$\boxed{\mathbf{L}(t) = \int_V \rho_0(\mathbf{X})\mathbf{V}(\mathbf{X},t)dV} \qquad \textbf{Linear Momentum (Material Form)} \qquad (3.6.4)$$

and the principle of linear momentum is now, using 3.1.31,

$$\frac{d}{dt}\int_V \rho_0(\mathbf{X})\mathbf{V}(\mathbf{X},t)dV = \int_V \rho_0 \frac{d\mathbf{V}}{dt} dV \equiv \mathbf{F}(t) \qquad (3.6.5)$$

The external forces **F** to be considered are those acting on the *current* configuration. Suppose that the surface force acting on a surface element ds in the current configuration is $d\mathbf{f}_{\text{surf}} = \mathbf{t}ds = \mathbf{T}dS$, where **t** and **T** are, respectively, the Cauchy traction vector and the PK1 traction vector (Eqns. 3.5.3-4). Also, just as the PK1 stress measures the actual force in the current configuration, but per unit surface area in the reference configuration, one can introduce the **reference body force B**: this is the actual body force acting in the current configuration, per unit volume in the reference configuration. Thus if the body force acting on a volume element dv in the current configuration is $d\mathbf{f}_{\text{body}}$, then

$$d\mathbf{f}_{\text{body}} = \mathbf{b}dv = \mathbf{B}dV \qquad (3.6.6)$$

The resultant force acting on the body is then

$$\mathbf{F}(t) = \int_S \mathbf{T}dS + \int_V \mathbf{B}dV, \qquad F_i = \int_S T_i\, dS + \int_V B_i\, dV \qquad (3.6.7)$$

Using Cauchy's law, $\mathbf{T} = \mathbf{PN}$, where **P** is the PK1 stress, and the divergence theorem 1.12.21, 3.6.5 and 3.6.7 lead to

$$\int_V [\text{Div}\mathbf{P} + \mathbf{B}]dV = \int_V \rho_0 \frac{d\mathbf{V}}{dt} dV \qquad (3.6.8)$$

and the corresponding local form is

$$\boxed{\text{Div}\,\mathbf{P} + \mathbf{B} = \rho_0 \frac{d\mathbf{V}}{dt}, \qquad \frac{\partial P_{ij}}{\partial X_j} + B_i = \rho_0 \frac{dV_i}{dt}}$$

Equations of Motion (Material Form) (3.6.9)

Derivation from the Spatial Form

The equations of motion can also be derived directly from the spatial equations. In order to do this, one must first show that $\mathrm{Div}\left(J\mathbf{F}^{-\mathrm{T}}\right)$ is zero. One finds that (using the divergence theorem, Nanson's formula 2.2.59 and the fact that $\mathrm{div}\mathbf{I} = 0$)

$$\begin{aligned}
&\int_V \mathrm{Div}\left(J\mathbf{F}^{-\mathrm{T}}\right) dV = \int_S J\mathbf{F}^{-\mathrm{T}}\mathbf{N}dS = \int_s \mathbf{n}ds = \int_s \mathbf{I}\mathbf{n}ds = \int_v \mathrm{div}\mathbf{I}\,dv = 0 \\
&\int_V \frac{\partial\left(JF_{ji}^{-1}\right)}{\partial X_j} dV = \int_S JF_{ji}^{-1}N_i dS = \int_s n_i ds = \int_s \delta_{ij} n_j ds = \int_v \frac{\partial \delta_{ij}}{\partial x_i} dv = 0
\end{aligned} \tag{3.6.10}$$

This result is known as the **Piola identity**. Thus, with the PK1 stress related to the Cauchy stress through 3.5.8, $\mathbf{P} = J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}$, and using identity 1.14.16c,

$$\begin{aligned}
\mathrm{Div}\mathbf{P} &= \mathrm{Div}\left(\boldsymbol{\sigma}\left(J\mathbf{F}^{-\mathrm{T}}\right)\right) \\
&= \boldsymbol{\sigma}\mathrm{Div}\left(J\mathbf{F}^{-\mathrm{T}}\right) + \mathrm{Grad}\boldsymbol{\sigma} : \left(J\mathbf{F}^{-\mathrm{T}}\right) \\
&= J\mathrm{Grad}\boldsymbol{\sigma} : \mathbf{F}^{-\mathrm{T}}
\end{aligned} \tag{3.6.11}$$

From 2.2.8c,

$$\mathrm{Div}\mathbf{P} = J\mathrm{div}\boldsymbol{\sigma} \tag{3.6.12}$$

Then, with $dv = JdV$ and 3.6.6, the equations of motion in the spatial form can now be transformed according to

$$\int_v \left[\mathrm{div}\boldsymbol{\sigma} + \mathbf{b}\right] dv = \int_v \rho\,\dot{\mathbf{v}}\,dv \quad \to \quad \int_V \left[\mathrm{Div}\mathbf{P} + \mathbf{B}\right] dV = \int_V \rho_0 \dot{\mathbf{V}}\,dV$$

as before.

3.6.3 Symmetry of the Cauchy Stress

It will now be shown that the principle of angular momentum leads to the requirement that the Cauchy stress tensor is symmetric. Applying Cauchy's law to 3.2.11,

$$\begin{aligned}
&\int_s \mathbf{r} \times (\boldsymbol{\sigma}\mathbf{n}) ds + \int_v \mathbf{r} \times \mathbf{b} dv = \frac{d}{dt}\int_v \mathbf{r} \times \rho\mathbf{v} dv \\
&\int_s \varepsilon_{ijk} x_j \sigma_{kl} n_l dS + \int_v \varepsilon_{ijk} x_j b_k\, dv = \frac{d}{dt}\int_v \varepsilon_{ijk} x_j \rho v_k\, dv
\end{aligned} \tag{3.6.13}$$

The surface integral can be converted into a volume integral using the divergence theorem. Using the index notation, and concentrating on the integrand of the resulting volume integral, one has, using 1.3.14 (the permutation symbol is a constant here, $\partial\varepsilon_{ijk} / \partial x_l = 0$),

$$\varepsilon_{ijk}\frac{\partial(x_j\sigma_{kl})}{\partial x_l}=\varepsilon_{ijk}\left\{x_j\frac{\partial\sigma_{kl}}{\partial x_l}+\sigma_{kl}\delta_{jl}\right\}=\varepsilon_{ijk}\left\{x_j\frac{\partial\sigma_{kl}}{\partial x_l}+\sigma_{kj}\right\}\equiv\mathbf{r}\times\operatorname{div}\boldsymbol{\sigma}+\mathsf{E}:\boldsymbol{\sigma}^{\mathrm{T}} \qquad (3.6.14)$$

where E is the third-order permutation tensor, Eqn. 1.9.6, $\mathsf{E}=\varepsilon_{ijk}(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k)$. Thus, with the Reynold's transport identity 3.1.31,

$$\int_v\{\mathbf{r}\times\operatorname{div}\boldsymbol{\sigma}+\mathsf{E}:\boldsymbol{\sigma}^{\mathrm{T}}\}dv+\int_v\mathbf{r}\times\mathbf{b}dv=\int_v\rho\frac{d}{dt}(\mathbf{r}\times\mathbf{v})dv \qquad (3.6.15)$$

The material derivative of this cross product is

$$\frac{d}{dt}(\mathbf{r}\times\mathbf{v})=\mathbf{r}\times\frac{d\mathbf{v}}{dt}+\frac{d\mathbf{r}}{dt}\times\mathbf{v}=\mathbf{r}\times\frac{d\mathbf{v}}{dt}+\mathbf{v}\times\mathbf{v}=\mathbf{r}\times\frac{d\mathbf{v}}{dt} \qquad (3.6.16)$$

and so

$$\int_v\mathsf{E}:\boldsymbol{\sigma}^{\mathrm{T}}dv+\int_v\mathbf{r}\times\left\{\operatorname{div}\boldsymbol{\sigma}+\mathbf{b}-\rho\frac{d\mathbf{v}}{dt}\right\}dv=0 \qquad (3.6.17)$$

From the equations of motion 2.6.2, the term inside the brackets is zero, so that

$$\mathsf{E}:\boldsymbol{\sigma}^{\mathrm{T}}=0,\qquad \varepsilon_{ijk}\sigma_{kj}=0 \qquad (3.6.18)$$

It follows, from expansion of this relation, that the matrix of stress components must be symmetric:

$$\boxed{\boldsymbol{\sigma}=\boldsymbol{\sigma}^{\mathrm{T}},\qquad \sigma_{ij}=\sigma_{ji}}\qquad \textbf{Symmetry of Stress} \qquad (3.6.19)$$

3.6.4 Consequences in the Material Form

Here, the consequences of 3.6.19 on the PK1 and PK2 stresses is examined. Using the result $\boldsymbol{\sigma}=\boldsymbol{\sigma}^{\mathrm{T}}$ and 3.5.8, $\boldsymbol{\sigma}=J^{-1}\mathbf{P}\mathbf{F}^{\mathrm{T}}$,

$$J^{-1}\mathbf{P}\mathbf{F}^{\mathrm{T}}=\left(J^{-1}\mathbf{P}\mathbf{F}^{\mathrm{T}}\right)^{\mathrm{T}}=J^{-1}\mathbf{F}\mathbf{P}^{\mathrm{T}} \qquad (3.6.20)$$

so that

$$\mathbf{P}\mathbf{F}^{\mathrm{T}}=\mathbf{F}\mathbf{P}^{\mathrm{T}},\qquad P_{ik}F_{jk}=F_{ik}P_{jk} \qquad (3.6.21)$$

These equations are trivial when $i=j$, not providing any constraint on **P**. On the other hand, when $i\neq j$ one has the three equations

$$
\begin{aligned}
P_{11}F_{21} + P_{12}F_{22} + P_{13}F_{23} &= F_{11}P_{21} + F_{12}P_{22} + F_{13}P_{23} \\
P_{11}F_{31} + P_{12}F_{32} + P_{13}F_{33} &= F_{11}P_{31} + F_{12}P_{32} + F_{13}P_{33} \\
P_{21}F_{31} + P_{22}F_{32} + P_{23}F_{33} &= F_{21}P_{31} + F_{22}P_{32} + F_{23}P_{33}
\end{aligned}
\qquad (3.6.22)
$$

Thus angular momentum considerations imposes these three constraints on the PK1 stress (as they imposed the three constraints $\sigma_{12} = \sigma_{21}$, $\sigma_{13} = \sigma_{31}$, $\sigma_{23} = \sigma_{32}$ on the Cauchy stress).

It has already been seen that a consequence of the symmetry of the Cauchy stress is the symmetry of the PK2 stress **S**; thus, formally, the symmetry of **S** is the result of the angular momentum principle.

3.7 Boundary Conditions and The Boundary Value Problem

In order to solve a mechanics problem, one must specify certain conditions around the boundary of the material under consideration. Such **boundary conditions** will be discussed here, together with the resulting **boundary value problem** (**BVP**). (see Part I, 3.5.1, for a discussion of stress boundary conditions.)

3.7.1 Boundary Conditions

There are two types of boundary condition, those on displacement and those on traction. Denote the body in the reference condition by B_0 and in the current configuration by *B*. Denote the boundary of the body in the reference configuration by S and in the current configuration by *s*, Fig. 3.7.1.

Displacement Boundary Conditions

The position of particles may be specified over some portion of the boundary in the current configuration. That is, $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X})$ is specified to be $\bar{\mathbf{x}}$ say, over some portion $s_{\mathbf{u}}$ of *s*, Fig. 3.7.1, which corresponds to the portion $S_{\mathbf{u}}$ of S. With $\mathbf{u}(\mathbf{x}) = \mathbf{x} - \mathbf{X}(\mathbf{x})$, or $\mathbf{U}(\mathbf{X}) = \mathbf{x}(\mathbf{X}) - \mathbf{X}$, this can be expressed as

$$\begin{aligned} \mathbf{u}(\mathbf{x}) &= \bar{\mathbf{u}}(\mathbf{x}), \quad && \mathbf{x} \in s_{\mathbf{u}} \\ \mathbf{U}(\mathbf{X}) &= \bar{\mathbf{U}}(\mathbf{X}), \quad && \mathbf{X} \in S_{\mathbf{u}} \end{aligned} \tag{3.7.1}$$

These are called **displacement boundary conditions**. The most commonly encountered displacement boundary condition is where some portion of the boundary is fixed, in which case $\bar{\mathbf{u}}(\mathbf{x}) = \mathbf{o}$.

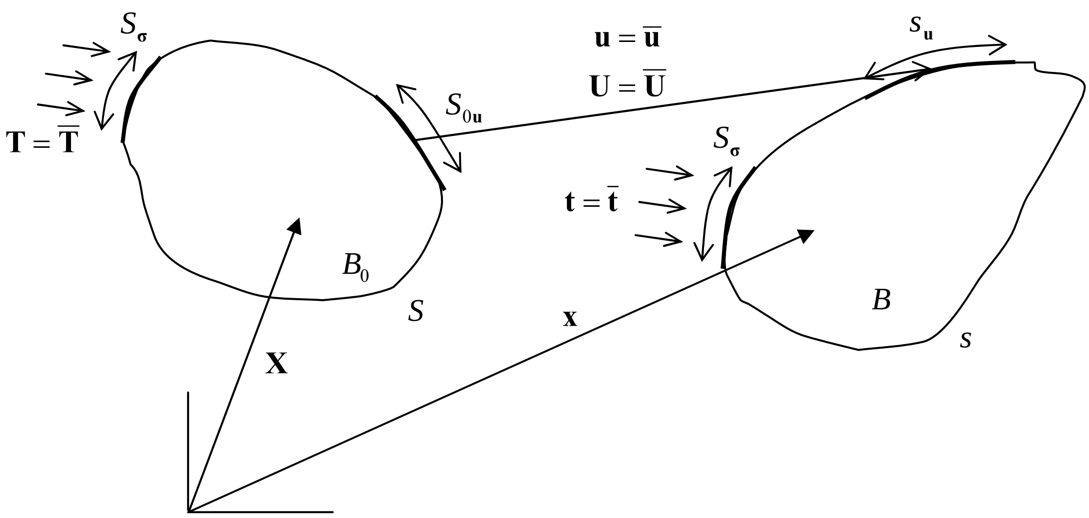


Figure 3.7.1: Boundary conditions

Traction Boundary Conditions

Traction $\mathbf{t} = \bar{\mathbf{t}}$ can be specified over a portion s_σ of the boundary, Fig. 3.7.1. These traction boundary conditions are related to the PK1 traction $\mathbf{T} = \overline{\mathbf{T}}$ over the corresponding surface S_σ in the reference configuration, through Eqns. 3.5.1-4,

$$\mathbf{T}dS = \mathbf{PN}dS = \mathbf{t}ds = \boldsymbol{\sigma}\mathbf{n}ds \tag{3.7.2}$$

One usually knows the position of the boundary S and the normal $\mathbf{N}(\mathbf{X})$ in the reference configuration. As deformation proceeds, the PK1 traction develops according to $\overline{\mathbf{T}} = \mathbf{PN}$ with, from 3.5.8, $\mathbf{P} = J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}$. The PK1 stress will in general depend on the motion $\mathbf{x}$ and the deformation gradient $\mathbf{F}$, so the traction boundary condition can be expressed in the general form

$$\overline{\mathbf{T}} = \overline{\mathbf{T}}(\mathbf{X}, \mathbf{x}, \mathbf{F}) \tag{3.7.3}$$

Example: Fluid Pressure

Consider the case of fluid pressure p around the boundary, $\bar{\mathbf{t}} = -p\mathbf{n}$, Fig. 3.7.2. The Cauchy traction $\bar{\mathbf{t}}$ depends through the normal $\mathbf{n}$ on the new position and geometry of the surface s_σ. Also, $\overline{\mathbf{T}} = -pJ\mathbf{F}^{-\mathrm{T}}\mathbf{N}$, which is of the general form 3.7.3.

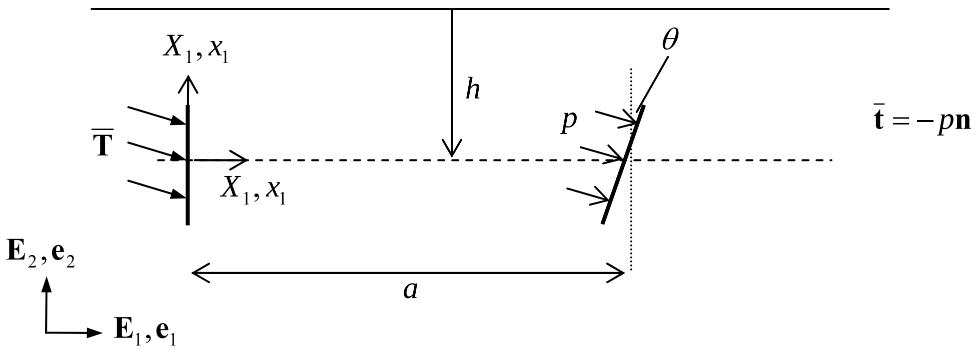


Figure 3.7.2: Fluid pressure on deforming material

Consider a material under water with part of its surface deforming as shown in Fig. 3.7.2. Referring to the figure, $\mathbf{N} = -\mathbf{E}_1$, $\mathbf{n} = -\cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2$, $\boldsymbol{\sigma} = -p\mathbf{I}$, $p = \rho g(h - x_2)$ and

$$\begin{aligned} x_1 &= X_1 + a + X_2 \tan\theta \\ x_2 &= X_2 \\ x_3 &= X_3 \end{aligned}, \quad \mathbf{F} = \begin{bmatrix} 1 & \tan\theta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad J = \det\mathbf{F} = 1$$

The traction vectors and PK1 stress are

$$\bar{\mathbf{t}} = -\rho g(h - x_2)\begin{bmatrix} -\cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \quad \overline{\mathbf{T}} = -\rho g(h - X_2)\begin{bmatrix} -1 \\ \tan\theta \\ 0 \end{bmatrix}, \quad \mathbf{P} = -\rho g(h - x_2)\begin{bmatrix} 1 & 0 & 0 \\ -\tan\theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with (note that $dS / ds = \cos\theta$) $|\bar{\mathbf{t}}| = p$ and $|\overline{\mathbf{T}}| = p / \cos\theta$. The traction vectors clearly depend on both position, and the deformation through θ. In this example, $\mathrm{grad}\mathbf{u} = \mathbf{F} - \mathbf{I} = \mathrm{Grad}\mathbf{U} = \mathbf{I} - \mathbf{F}^{-1} = \tan\theta \mathbf{e}_1 \otimes \mathbf{e}_2$ and

$$\theta(\mathrm{grad}\mathbf{u}) = \arctan\|\mathrm{grad}\mathbf{u}\| = \arctan\sqrt{\mathrm{grad}\mathbf{u} : \mathrm{grad}\mathbf{u}}$$

■

Dead Loading

A special case of loading is that of **dead loading**, where

$$\overline{\mathbf{T}} = \overline{\mathbf{T}}(\mathbf{X}) \tag{3.7.4}$$

Here, the PK1 stress on the boundary does not change with the deformation and an initially normal traction will not remain so as deformation proceeds.

For example, if one considers again the geometry of Fig. 3.7.2, this time take

$$\overline{\mathbf{T}}(\mathbf{X}) = \mathbf{P}\mathbf{N} = -p\mathbf{N} = \rho g(h - X_2)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{P}(\mathbf{X}) = -\rho g(h - X_2)\mathbf{I}$$

Then

$$\bar{\mathbf{t}}(\mathbf{x}, \theta) = \cos\theta \rho g(h - x_2)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \boldsymbol{\sigma}(\mathbf{x}, \theta) = -\rho g(h - x_2)\begin{bmatrix} 1 & 0 & 0 \\ \tan\theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3.7.2 The Boundary Value Problem

The equations of motion 3.6.2, 3.6.9, are a set of three differential equations. In the solution of any problem, one would have to supplement these equations with others, for example a constitutive equation expressing a relationship between the stress and the kinematic variables (see Part IV). This constitutive relation will typically relate the stress to the strains, or rates of strain, for example $\boldsymbol{\sigma} = f(\mathbf{e}, \mathbf{d})$. Suppose then that the stresses are known in terms of the strains and hence the displacements $\mathbf{u}$. The equations of motion are then a set of three second order differential equations in the three unknowns u_i (assuming that the body force $\mathbf{b}$ is a prescribed function of the problem). They need to be subjected to certain boundary and initial conditions.

Assume that the boundary conditions are such that the displacements are specified over that part of the surface $s_{\mathbf{u}}$ and tractions are specified over that part $s_{\boldsymbol{\sigma}}$, with the total surface $s = s_{\mathbf{u}} + s_{\boldsymbol{\sigma}}$, with $s_{\mathbf{u}} \cap s_{\boldsymbol{\sigma}} = 0$ [1]. Thus

$$\begin{aligned} \mathbf{t} = \boldsymbol{\sigma}\mathbf{n} = \bar{\mathbf{t}}, \qquad & \text{on } s_\sigma \\ \mathbf{u} = \bar{\mathbf{u}}, \qquad & \text{on } s_u \end{aligned} \qquad \textbf{Boundary Conditions} \qquad (3.7.5)$$

where the overbar signifies quantities which are prescribed. Initial conditions are also required for the displacement and velocity, so that

$$\begin{aligned} \mathbf{u}(\mathbf{x},\mathbf{t}) = \mathbf{u}_0(\mathbf{x}), \qquad & \text{at } t = 0 \\ \dot{\mathbf{u}}(\mathbf{x},\mathbf{t}) = \dot{\mathbf{u}}_0(\mathbf{x}), \qquad & \text{at } t = 0 \end{aligned} \qquad \textbf{Initial Conditions} \qquad (3.7.6)$$

and it is usually taken that $\mathbf{x} = \mathbf{X}$ at $t = 0$. Comparing 3.7.5 and 3.7.6, one also requires that $\mathbf{u}_0 = \bar{\mathbf{u}}$, $\dot{\mathbf{u}}_0 = \dot{\bar{\mathbf{u}}}$ over $s_{\mathbf{u}}$, so that the boundary and initial conditions are compatible.

These equations together, the differential equations of motion and the boundary and initial conditions, are called the **strong form** of the initial boundary value problem (BVP):

$$\boxed{\begin{array}{ll} \multicolumn{2}{c}{\mathrm{div}\boldsymbol{\sigma} + \mathbf{b} = \rho\dot{\mathbf{v}} = \rho\ddot{\mathbf{u}}} \\ \mathbf{t} = \boldsymbol{\sigma}\mathbf{n} = \bar{\mathbf{t}}, & \text{on } s_\sigma \\ \mathbf{u} = \bar{\mathbf{u}}, & \text{on } s_u \\ \mathbf{u}(\mathbf{x},\mathbf{t}) = \mathbf{u}_0(\mathbf{x}), & \text{at } t = 0 \\ \dot{\mathbf{u}}(\mathbf{x},\mathbf{t}) = \dot{\mathbf{u}}_0(\mathbf{x}), & \text{at } t = 0 \end{array}} \qquad \textbf{Strong form of the Initial BVP} \qquad (3.7.7)$$

When the problem is quasi-static, so the accelerations can be neglected, the equations of motion reduce to the equations of equilibrium 3.6.3. In that case one does not need initial conditions and one has a boundary value problem involving 3.7.5 only.

It is only in certain special cases and in certain simple problems that an exact solution can be obtained to these equations. An alternative solution strategy is to convert these equations into what is known as the **weak form**. The weak form, which is in the form of integrals rather than differential equations, can then be solved approximately using a numerical technique, for example the Finite Element Method[2]. The weak form is discussed in §3.9.

---

[1] It is possible to specify both traction and displacement over the same portion of the boundary, but not the same components. For example, if one specified $\mathbf{t} = t_1\mathbf{e}_1$ on a boundary, one could also specify $\mathbf{u} = u_2\mathbf{e}_2$, but not $\mathbf{u} = u_1\mathbf{e}_1$. In that case, one could imagine the boundary to consist of two separate boundaries, one with conditions with respect to $\mathbf{e}_1$ and one with respect to $\mathbf{e}_2$, and still write $s_{\mathbf{u}} \cap s_{\boldsymbol{\sigma}} = 0$.

[2] Further, it is often easier to prove results regarding the uniqueness and stability of solutions to the problem when it is cast in the weak form

In the material form, the boundary conditions are

$$\begin{aligned} &\mathbf{T} = \mathbf{P}\mathbf{N} = \overline{\mathbf{T}}, && \text{on } S_\sigma \\ &\mathbf{U} = \overline{\mathbf{U}}, && \text{on } S_u \end{aligned} \qquad \textbf{Boundary Conditions} \qquad (3.7.8)$$

and the initial conditions are

$$\begin{aligned} &\mathbf{U}(\mathbf{X},\mathbf{t}) = \mathbf{U}_0(\mathbf{X}), && \text{at } t = 0 \\ &\dot{\mathbf{U}}(\mathbf{X},\mathbf{t}) = \dot{\mathbf{U}}_0(\mathbf{X}), && \text{at } t = 0 \end{aligned} \qquad \textbf{Initial Conditions} \qquad (3.7.9)$$

and the initial vale problem is

$$\boxed{\begin{array}{ll} \multicolumn{2}{c}{\mathrm{Div}\mathbf{P} + \mathbf{B} = \rho_0 \dot{\mathbf{V}} = \rho \ddot{\mathbf{U}}} \\ \mathbf{T} = \mathbf{P}\mathbf{N} = \overline{\mathbf{T}}, & \text{on } S_\sigma \\ \mathbf{U} = \overline{\mathbf{U}}, & \text{on } S_u \\ \mathbf{U}(\mathbf{X},\mathbf{t}) = \mathbf{U}_0(\mathbf{X}), & \text{at } t = 0 \\ \dot{\mathbf{U}}(\mathbf{X},\mathbf{t}) = \dot{\mathbf{U}}_0(\mathbf{X}), & \text{at } t = 0 \end{array}} \qquad \textbf{Strong form of the Initial BVP} \qquad (3.7.10)$$

3.8 Balance of Mechanical Energy

3.8.1 The Balance of Mechanical Energy

First, from Part I, Chapter 5, recall work and kinetic energy are related through

$$W_{\text{ext}} + W_{\text{int}} = \Delta K \tag{3.8.1}$$

where W_{ext} is the work of the external forces and W_{int} is the work of the internal forces. The *rate* form is

$$P_{\text{ext}} + P_{\text{int}} = \dot{K} \tag{3.8.2}$$

where the external and internal *powers* and rate of change of kinetic energy are

$$P_{\text{ext}} = \frac{d}{dt}W_{\text{ext}}, \quad P_{\text{int}} = \frac{d}{dt}W_{\text{int}}, \quad \dot{K} = \frac{d}{dt}\Delta K \tag{3.8.3}$$

This expresses the *mechanical* energy balance for a material. Eqn. 3.8.2 is equivalent to the equations of motion (see below).

The total external force acting on the material is given by 3.2.6:

$$\mathbf{F}_{\text{ext}} = \int_s \mathbf{t}\, ds + \int_v \mathbf{b}\, dv \tag{3.8.4}$$

The increment in work done dW when an element subjected to a body force (per unit volume) $\mathbf{b}$ undergoes a displacement $d\mathbf{u}$ is $\mathbf{b} \cdot d\mathbf{u}\, dv$. The rate of working is $dP = \mathbf{b} \cdot (d\mathbf{u} / dt) dv$. Thus, and similarly for the traction, the power of the external forces is

$$P_{\text{ext}} = \int_s \mathbf{t} \cdot \mathbf{v}\, ds + \int_v \mathbf{b} \cdot \mathbf{v}\, dv \tag{3.8.5}$$

where $\mathbf{v}$ is the velocity. Also, the total kinetic energy of the matter in the volume is

$$K = \int_v \tfrac{1}{2} \rho \mathbf{v} \cdot \mathbf{v} dv \tag{3.8.6}$$

Using Reynold's transport theorem,

$$\frac{d}{dt}K = \int_v \tfrac{1}{2} \rho \frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) dv = \int_v \rho \mathbf{v} \cdot \frac{d\mathbf{v}}{dt} dv \tag{3.8.7}$$

Thus the expression 3.8.2 becomes

$$\int_s \mathbf{t}\cdot\mathbf{v}\,ds + \int_v \mathbf{b}\cdot\mathbf{v}\,dv + P_{\text{int}} = \int_v \rho\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\,dv \tag{3.8.8}$$

power of surface forces | power of body forces | power of internal forces | rate of change of kinetic energy

It can be seen that some of the power exerted by the external forces alters the kinetic energy of the material and the remainder changes its internal energy state.

Conservative Force System

In the special case where the internal forces are conservative, that is, no energy is dissipated as heat, but all energy is stored as internal energy, one can express the power of the internal forces in terms of a potential function u (see Part I, §5.1), and rewrite this equation as

$$\int_s \mathbf{t}\cdot\mathbf{v}\,ds + \int_v \mathbf{b}\cdot\mathbf{v}\,dv = \int_v \rho\frac{du}{dt}\,dv + \int_v \rho\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\,dv \tag{3.8.9}$$

rate of change of internal energy

Here, the rate of change of the internal energy has been written in the form

$$\frac{d}{dt}U = \frac{d}{dt}\int_v \rho u dv = \int_v \rho\frac{du}{dt}dv \tag{3.8.10}$$

where u is the internal energy per unit mass, or the **specific internal energy**.

3.8.2 The Stress Power

To express the power of the internal forces P_{int} in terms of stresses and strain-rates, first re-write the rate of change of kinetic energy using the equations of motion,

$$\frac{d}{dt}K = \int_v \rho\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}dv = \int_v \mathbf{v}\cdot\left(\operatorname{div}\boldsymbol{\sigma} + \mathbf{b}\right)dv \tag{3.8.11}$$

Also, using the product rule of differentiation,

$$\mathbf{v}\cdot\operatorname{div}\boldsymbol{\sigma} = \operatorname{div}(\mathbf{v}\boldsymbol{\sigma}) - \boldsymbol{\sigma}:\mathbf{l}, \qquad v_i\frac{\partial\sigma_{ij}}{\partial x_j} = \frac{\partial(v_i\sigma_{ij})}{\partial x_j} - \sigma_{ij}\frac{\partial v_i}{\partial x_j} \tag{3.8.12}$$

where $\mathbf{l}$ is the spatial velocity gradient, $l_{ij} = \partial v_i / \partial x_j$. Decomposing $\mathbf{l}$ into its symmetric part $\mathbf{d}$, the rate of deformation, and its antisymmetric part $\mathbf{w}$, the spin tensor, gives

$$\boldsymbol{\sigma} : \mathbf{l} = \boldsymbol{\sigma} : \mathbf{d} + \boldsymbol{\sigma} : \mathbf{w} = \boldsymbol{\sigma} : \mathbf{d}, \qquad \sigma_{ij} \frac{\partial v_i}{\partial x_j} = \sigma_{ij} \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \tag{3.8.13}$$

since the double contraction of any symmetric tensor ($\boldsymbol{\sigma}$) with any skew-symmetric tensor ($\mathbf{w}$) is zero, 1.10.31c. Also, using Cauchy's law and the divergence theorem 1.14.21,

$$\begin{gathered} \int_s \mathbf{t} \cdot \mathbf{v}\, ds = \int_s \boldsymbol{\sigma}\mathbf{n} \cdot \mathbf{v}\, ds = \int_s (\mathbf{v}\boldsymbol{\sigma}) \cdot \mathbf{n}\, ds = \int_v \operatorname{div}(\mathbf{v}\boldsymbol{\sigma}) dv \\ \int_s t_i v_i \, ds = \int_s \sigma_{ik} n_k v_i \, ds = \int_v \frac{\partial(\sigma_{ik} v_i)}{\partial x_k} dv \end{gathered} \tag{3.8.14}$$

Thus, finally, from Eqn. 3.8.8,

$$\boxed{P_{\text{int}} = -\int_v \boldsymbol{\sigma} : \mathbf{d}\, dv, \qquad P_{\text{int}} = -\int_v \sigma_{ij} \left\{ \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \right\} dv} \quad \textbf{Stress Power} \tag{3.8.15}$$

The term $\boldsymbol{\sigma} : \mathbf{d}$ is called the **stress power**; the stress power is the (negative of the) rate of working of the internal forces, per unit volume. The complete equation for the conservation of mechanical energy is then

$$\boxed{\int_s \mathbf{t} \cdot \mathbf{v}\, ds + \int_v \mathbf{b} \cdot \mathbf{v}\, dv = \int_v \boldsymbol{\sigma} : \mathbf{d}\, dv + \int_v \rho \mathbf{v} \cdot \frac{d\mathbf{v}}{dt} dv} \quad \textbf{Mechanical Energy Balance} \tag{3.8.16}$$

The stress power is that part of the externally supplied power which is not converted into kinetic energy; it is converted into heat and a change in internal energy.

Note that, as with the law of conservation of mechanical energy for a particle, this equation does not express a separate law of continuum mechanics; it is merely a re-arrangement of the equations of motion (see below), which themselves follows from the principle of linear momentum (Newtons second law).

Conservative Force System

If the internal forces are conservative, one has

$$\int_v \boldsymbol{\sigma} : \mathbf{d}\, dv = \frac{d}{dt} U = \int_v \rho \frac{du}{dt} dv \tag{3.8.17}$$

or, in local form,

$$\boxed{\boldsymbol{\sigma} : \mathbf{d} = \rho \frac{du}{dt}} \quad \textbf{Mechanical Energy Balance (Conservative System)} \qquad (3.8.18)$$

This is the local form of the energy equation for the case of a purely mechanical conservative process.

3.8.3 Derivation from the Equations of Motion

As mentioned, the conservation of mechanical energy equation can be derived directly from the equations of motion. The derivation is similar to that used above (where the mechanical energy equations were used to derive an expression for the stress power using the equations of motion). One has, multiplying the equations of motion by **v** and integrating,

$$\begin{aligned}\int_v \rho \mathbf{v} \cdot \frac{d\mathbf{v}}{dt} dv = \int_v \mathbf{v} \cdot (\operatorname{div} \boldsymbol{\sigma} + \mathbf{b}) dv &= \int_v \{\operatorname{div}(\mathbf{v} \cdot \boldsymbol{\sigma}) - \boldsymbol{\sigma} : \mathbf{l} + \mathbf{v} \cdot \mathbf{b}\} dv \\ &= \int_v \{\operatorname{div}(\mathbf{v} \cdot \boldsymbol{\sigma}) - \boldsymbol{\sigma} : \mathbf{d} + \mathbf{v} \cdot \mathbf{b}\} dv \\ &= \int_v - \boldsymbol{\sigma} : \mathbf{d} dv + \int_s \mathbf{t} \cdot \mathbf{v} ds + \int_v \mathbf{v} \cdot \mathbf{b} dv\end{aligned} \qquad (3.8.19)$$

3.8.4 Stress Power and the Continuum Element

In the above, the stress power was derived using a global (integral) form of the equations. The stress power can also be deduced by considering a differential mass element. For example, consider such an element whose boundary particles are moving with velocity **v** and whose boundary is subjected to stresses $\boldsymbol{\sigma}$, Fig. 3.8.1.

Consider first the components of force and velocity acting in the x_1 direction. The external forces act on the six sides. On three of them (the ones that can be seen in the illustration) the stress and velocity act in the same direction, so the power is positive; on the other three they act in opposite directions, so there the power is negative.

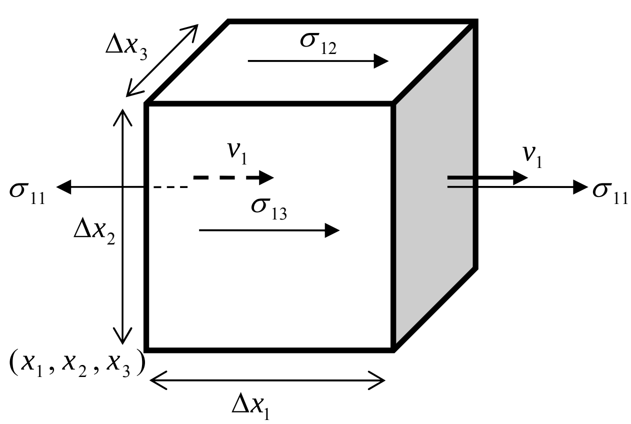


Figure 3.8.1: A differential mass element subjected to stresses

As usual (see §1.6.6), the element is assumed to be small enough so that the product of stress and velocity varies linearly over the element, so that the average of this product over an element face can be taken to be representative of the power of the surface forces on that element. The power of the external surface forces acting on the three faces to the front is then

$$\begin{aligned} P_{\text{surf}} = \Delta x_2 \Delta x_3 (\sigma_{11} v_1)_{x_1+\Delta x_1, x_2+\frac{1}{2}\Delta x_2, x_3+\frac{1}{2}\Delta x_3} &+ \Delta x_1 \Delta x_3 (\sigma_{12} v_1)_{x_1+\frac{1}{2}\Delta x_1, x_2+\Delta x_2, x_3+\frac{1}{2}\Delta x_3} \\ &+ \Delta x_1 \Delta x_2 (\sigma_{13} v_1)_{x_1+\frac{1}{2}\Delta x_1, x_2+\frac{1}{2}\Delta x_2, x_3+\Delta x_3} \end{aligned} \tag{3.8.20}$$

Using a Taylor's series expansion, and neglecting higher order terms, then leads to

$$\begin{aligned} P_{\text{surf}} \approx \Delta x_2 \Delta x_3 &\left\{ (\sigma_{11} v_1)_{x_1,x_2,x_3} + \Delta x_1 \frac{\partial(\sigma_{11} v_1)}{\partial x_1} + \tfrac{1}{2}\Delta x_2 \frac{\partial(\sigma_{11} v_1)}{\partial x_2} + \tfrac{1}{2}\Delta x_3 \frac{\partial(\sigma_{11} v_1)}{\partial x_3} \right\} \\ + \Delta x_1 \Delta x_3 &\left\{ (\sigma_{12} v_1)_{x_1,x_2,x_3} + \tfrac{1}{2}\Delta x_1 \frac{\partial(\sigma_{12} v_1)}{\partial x_1} + \Delta x_2 \frac{\partial(\sigma_{12} v_1)}{\partial x_2} + \tfrac{1}{2}\Delta x_3 \frac{\partial(\sigma_{12} v_1)}{\partial x_3} \right\} \\ + \Delta x_1 \Delta x_2 &\left\{ (\sigma_{13} v_1)_{x_1,x_2,x_3} + \tfrac{1}{2}\Delta x_1 \frac{\partial(\sigma_{13} v_1)}{\partial x_1} + \tfrac{1}{2}\Delta x_2 \frac{\partial(\sigma_{13} v_1)}{\partial x_2} + \Delta x_3 \frac{\partial(\sigma_{13} v_1)}{\partial x_3} \right\} \end{aligned} \tag{3.8.21}$$

The net power *per unit volume* (subtracting the power of the stresses on the other three surfaces and dividing through by the volume) is then

$$P_{\text{surf}} = \frac{\partial(\sigma_{11} v_1)}{\partial x_1} + \frac{\partial(\sigma_{12} v_1)}{\partial x_2} + \frac{\partial(\sigma_{13} v_1)}{\partial x_3} = \frac{\partial(\sigma_{1j} v_1)}{\partial x_j} \tag{3.8.22}$$

Assume the body force **b** to act at the centre of the element. Neglecting higher order terms which vanish as the element size is allowed to shrink towards zero, the power of the body force in the x_1 direction, per unit volume, is simply $b_1 v_1$.

The total power of the external forces is then (including the other two components of stress and velocity), using the equations of motion,

$$\begin{aligned} P_{\text{ext}} &= \frac{\partial(\sigma_{ij} v_i)}{\partial x_j} + b_i v_i \\ &= \sigma_{ij} \frac{\partial v_i}{\partial x_j} + \frac{\partial \sigma_{ij}}{\partial x_j} v_i + b_i v_i \\ &= \sigma_{ij} \frac{\partial v_i}{\partial x_j} + \left\{ -b_i + \rho \frac{dv_i}{dt} \right\} v_i + b_i v_i \\ &= \sigma_{ij} \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + \rho \frac{1}{2} \frac{d(v_i v_i)}{dt} \end{aligned} \qquad \begin{aligned} P_{\text{ext}} &= \operatorname{div}(\boldsymbol{\sigma}^{\mathrm{T}} \mathbf{v}) + \mathbf{b} \cdot \mathbf{v} \\ &= \boldsymbol{\sigma} : \mathbf{l} + \operatorname{div}\boldsymbol{\sigma} \cdot \mathbf{v} + \mathbf{b} \cdot \mathbf{v} \\ &= \boldsymbol{\sigma} : \mathbf{l} + \left\{ -\mathbf{b} + \rho \frac{d\mathbf{v}}{dt} \right\} \cdot \mathbf{v} + \mathbf{b} \cdot \mathbf{v} \\ &= \boldsymbol{\sigma} : \mathbf{d} + \rho \frac{1}{2} \frac{d(\mathbf{v} \cdot \mathbf{v})}{dt} \end{aligned} \tag{3.8.23}$$

which again equals the stress power term plus the change in kinetic energy.

The power of the internal forces is $-\boldsymbol{\sigma} : \mathbf{d}$, a result of the forces acting *inside* the differential element, *reacting* to the applied forces $\boldsymbol{\sigma}$ and $\mathbf{b}$.

3.8.5 The Balance of Mechanical Energy (Material form)

The material form of the power of the external forces is written as a function of the PK1 traction $\mathbf{T}$ and the reference body force $\mathbf{B}$, 3.6.7, and the kinetic energy as a function of the velocity $\mathbf{V}(\mathbf{X})$:

$$\int_S \mathbf{T} \cdot \mathbf{V}\, dS + \int_V \mathbf{B} \cdot \mathbf{V}\, dV + P_{\text{int}} = \int_V \rho_0 \mathbf{V} \cdot \frac{d\mathbf{V}}{dt} dV \tag{3.8.24}$$

Next, using the identities 2.5.4, $\dot{\mathbf{F}} = \mathbf{lF}$ and 1.10.3h, $\mathbf{A} : (\mathbf{BC}) = \left(\mathbf{AC}^{\mathrm{T}}\right) : \mathbf{B}$, gives

$$\boldsymbol{\sigma} : \mathbf{d} = \boldsymbol{\sigma} : \mathbf{l} - \boldsymbol{\sigma} : \mathbf{w} = \boldsymbol{\sigma} : \mathbf{l} = \boldsymbol{\sigma} : \left(\dot{\mathbf{F}}\mathbf{F}^{-1}\right) = \left(\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right) : \dot{\mathbf{F}}, \tag{3.8.25}$$

and so

$$\begin{aligned} \int_v \boldsymbol{\sigma} : \mathbf{d}\, dv &= \int_v \left(\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right) : \dot{\mathbf{F}}\, dv = \int_V \left(\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right) : \dot{\mathbf{F}}\, J dV \\ &= \int_V \mathbf{P} : \dot{\mathbf{F}}\, dV \end{aligned} \tag{3.8.26}$$

and

$$\boxed{\int_S \mathbf{T} \cdot \mathbf{V}\, dS + \int_V \mathbf{B} \cdot \mathbf{V}\, dV = \int_V \mathbf{P} : \dot{\mathbf{F}}\, dV + \int_V \rho_0 \mathbf{V} \cdot \frac{d\mathbf{V}}{dt} dV}$$

Mechanical Energy Balance (Material Form) (3.8.27)

For a conservative system, this can be written in terms of the internal energy

$$\int_S \mathbf{T} \cdot \mathbf{V}\, dS + \int_V \mathbf{B} \cdot \mathbf{V}\, dV = \int_V \rho_0 \frac{du}{dt} dV + \int_V \rho_0 \mathbf{V} \cdot \frac{d\mathbf{V}}{dt} dV \tag{3.8.28}$$

3.8.6 Work Conjugate Variables

Since the stress power is the double contraction of the Cauchy stress and rate-of-deformation, one says that the Cauchy stress and rate of deformation are **work conjugate** (or **power conjugate** or **energy conjugate**). Similarly, from 3.8.26, the PK1 stress $\mathbf{P}$ is power conjugate to $\dot{\mathbf{F}}$. It can also be shown that the PK2 stress $\mathbf{S}$ is power conjugate to the rate of Euler-Lagrange strain, $\dot{\mathbf{E}}$ (and hence also the right Cauchy-Green strain) {▲Problem 1} :

$$J\boldsymbol{\sigma} : \mathbf{d} = \mathbf{P} : \dot{\mathbf{F}} = \mathbf{S} : \dot{\mathbf{E}} = \mathbf{S} : \tfrac{1}{2}\dot{\mathbf{C}} \tag{3.8.29}$$

Note that, for conservative systems, these quantities represent the rate of change of internal energy per unit *reference* volume.

Using the polar decomposition and the relation $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$,

$$\begin{aligned}\dot{\mathbf{F}} &= \dot{\mathbf{R}}\mathbf{U} + \mathbf{R}\dot{\mathbf{U}} \\ &= \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}}\mathbf{R}\mathbf{U} + \mathbf{R}\dot{\mathbf{U}} \\ &= \boldsymbol{\Omega}_{\mathbf{R}}\mathbf{F} + \mathbf{R}\dot{\mathbf{U}}\end{aligned} \tag{3.8.30}$$

where $\boldsymbol{\Omega}_{\mathbf{R}}$ is the angular velocity tensor 2.6.1. Then, using 1.11.3h, 1.10.31c, and the definitions 3.5.8, 3.5.12 and 3.5.18,

$$\begin{aligned}\mathbf{P} : \dot{\mathbf{F}} &= \mathbf{P} : \boldsymbol{\Omega}_{\mathbf{R}}\mathbf{F} + \mathbf{P} : \mathbf{R}\dot{\mathbf{U}} \\ &= \mathbf{P}\mathbf{F}^{\mathrm{T}} : \boldsymbol{\Omega}_{\mathbf{R}} + \mathbf{R}^{\mathrm{T}}\mathbf{P} : \dot{\mathbf{U}} \\ &= \boldsymbol{\tau} : \boldsymbol{\Omega}_{\mathbf{R}} + \mathbf{R}^{\mathrm{T}}\mathbf{P} : \dot{\mathbf{U}} \\ &= \mathbf{T}_{\mathrm{B}} : \dot{\mathbf{U}}\end{aligned} \tag{3.8.31}$$

so that the Biot stress is power conjugate to the right stretch tensor. Since $\mathbf{U}$ is symmetric, $\mathbf{P} : \dot{\mathbf{F}} = \mathrm{sym}\mathbf{T}_{\mathrm{B}} : \dot{\mathbf{U}}$. Also, the Biot stress is conjugate to the Biot strain tensor $\overline{\mathbf{B}} = \mathbf{U} - \mathbf{I}$ introduced in §2.2.5.

From 3.5.14 and 1.10.3h,

$$\boldsymbol{\sigma} : \mathbf{d} = \mathbf{R}\hat{\boldsymbol{\sigma}}\mathbf{R}^{\mathrm{T}} : \mathbf{d} = \hat{\boldsymbol{\sigma}} : \hat{\mathbf{d}} \tag{3.8.32}$$

so that the corotational stress is power conjugate to the **rotated deformation rate**, defined by

$$\hat{\mathbf{d}} = \mathbf{R}^{\mathrm{T}}\mathbf{d}\mathbf{R} \tag{3.8.33}$$

Pull Back and Push Forward

From 2.12.12-13, the double contraction of two tensors can be expressed as push-forwards and pull-backs of those tensors. For example, the stress power (per unit reference volume) in the material description is $\mathbf{S} : \dot{\mathbf{E}}$. Then, using 3.5.13, 2.12.9a and 2.5.18b, $\dot{\mathbf{E}} = \mathbf{F}^{\mathrm{T}}\mathbf{d}\mathbf{F}$,

$$\mathbf{S} : \dot{\mathbf{E}} = \chi_*(\mathbf{S})^{\#} : \chi_*(\dot{\mathbf{E}})^{b} = \boldsymbol{\tau} : \mathbf{d} = J\boldsymbol{\sigma} : \mathbf{d} \tag{3.8.34}$$

This means that the material and spatial descriptions of the internal power can be transformed into each other using push-forward and pull-back operations.

Similarly, pulling back the corotational stress and rotated deformation rate to the intermediate configuration of Fig. 2.10.8, using 2.12.13, 2.12.27,

$$\boldsymbol{\sigma} : \mathbf{d} = \chi_*^{-1}(\boldsymbol{\sigma})^{\#}{}_{\mathbf{R}(\mathbf{g})} : \chi_*^{-1}(\mathbf{d})^{b}{}_{\mathbf{R}(\mathbf{g})} = \hat{\boldsymbol{\sigma}} : \hat{\mathbf{d}} \tag{3.8.35}$$

The stress power in terms of spatial tensors can also be expressed as a derivative of a tensor, using the Lie derivative. From 2.12.42, the Lie derivative of the Euler-Almansi strain is the rate of deformation and hence (note that there is no universal function whose derivative is **d**), so

$$J\boldsymbol{\sigma} : \mathbf{d} = J\boldsymbol{\sigma} : \mathrm{L}_{\mathrm{v}}^{b}\mathbf{e} \tag{3.8.36}$$

3.8.7 Problems

1. Show that the rate of internal energy per unit reference volume $J\boldsymbol{\sigma} : \mathbf{d}$ is equivalent to $\mathbf{S} : \dot{\mathbf{E}}$ (without using push-forwards/pull-backs).

3.9 The Principles of Virtual Work and Power

The principle of virtual work was introduced and discussed in Part I, §5.5. As mentioned there, it is yet another re-statement of the work – energy principle, only it is couched in terms of virtual displacements, and the principle of virtual power to be introduced below is an equivalent statement based on **virtual velocities**.

On the one hand, the principle of virtual work/power can be regarded as the fundamental law of dynamics for a continuum, and from it can be derived the equations of motion. On the other hand, one can regard the principle of linear momentum as the fundamental law, derive the equations of motion, and hence derive the principle of virtual work.

3.9.1 Overview of The Principle of Virtual Work

Consider a material under the action of external forces: body forces **b** and tractions **t**. The body undergoes a displacement $\mathbf{u}(\mathbf{x})$ due to these forces and now occupies its current configuration, Fig. 3.9.1. The problem is to find this displacement function $\mathbf{u}$.

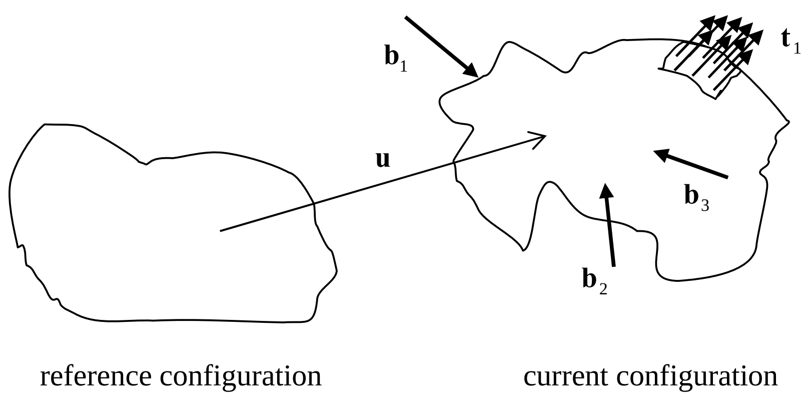


Figure 3.9.1: a material displacing to its current configuration under the action of body forces and surface forces

Imagine the material to undergo a small displacement $\delta\mathbf{u}$ from the *current configuration*, Fig. 3.9.2, $\delta\mathbf{u}$ not necessarily constant throughout the body; $\delta\mathbf{u}$ is a virtual displacement, meaning that it is an *imaginary* displacement, and in no way is it related to the applied external forces – *it does not actually occur physically*.

As each material particle moves through these virtual displacements, the external forces do virtual work δW. If the force $\mathbf{b}$ acts at position $\mathbf{x}$ and this point undergoes a virtual displacement $\delta u(\mathbf{x})$, the virtual work is $\delta W = \mathbf{b} \cdot \delta\mathbf{u}\,\Delta v$. Similarly for the surface tractions, and the total external virtual work is

$$\boxed{\delta W_{\text{ext}} = \int_v \mathbf{b} \cdot \delta\mathbf{u}\, dv + \int_s \mathbf{t} \cdot \delta\mathbf{u}\, ds = 0} \qquad \textbf{External Virtual Work} \qquad (3.9.1)$$

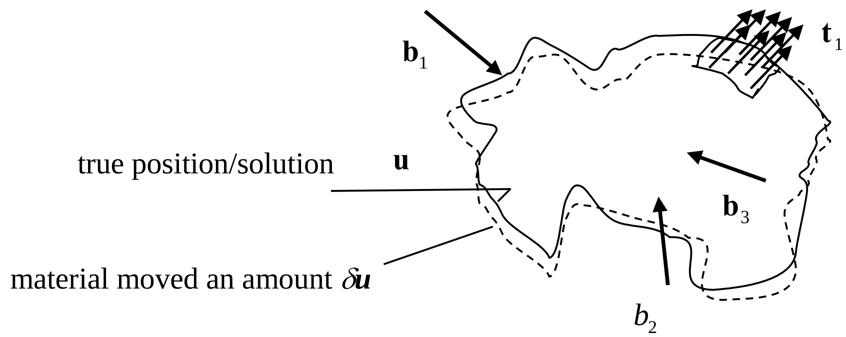


Figure 3.9.2: a virtual displacement field applied to a material in the current configuration

There is also an internal virtual work W_{int} due to the internal forces as they move through virtual displacements and a virtual kinetic energy δK. The principle of virtual work then says that

$$\delta W_{\text{ext}} + \delta W_{\text{int}} = \delta K \tag{3.9.2}$$

And this equation is then solved for the actual displacement **u**. Expressions for the internal virtual work and virtual kinetic energy will be derived below.

3.9.2 Derivation of The Principle of Virtual Work

As mentioned above, one can simply write down the principle of virtual work, regarding it as the fundamental principle of mechanics, and then from it derive the equations of motion. This will be done further below. To begin, though, the starting point will be the equations of motion, and from it will be derived the principle of virtual work.

Kinematically and Statically Admissible Fields

A **kinematically admissible displacement field** is defined to be one which satisfies the displacement boundary condition 3.7.7c, $\mathbf{u} = \overline{\mathbf{u}}$ on s_u (see Part I, §5.5.1). Such a displacement field would induce some stress field within the body, but this resulting stress field might not satisfy the equations of motion 3.7.7a. In other words, it might not be the actual displacement field, but it does not violate the boundary conditions.

A **statically admissible stress field** is one which satisfies the equations of motion 3.7.7a and the traction boundary conditions 3.7.7b, $\mathbf{t} = \boldsymbol{\sigma}\mathbf{n} = \overline{\mathbf{t}}$, on s_σ. Again, it might not be the actual stress field, since it is not specified how this stress field should be related to the actual displacement field.

Derivation from the Equations of Motion (Spatial Form)

Let $\boldsymbol{\sigma}$ be a statically admissible stress field corresponding to a kinematically admissible displacement field $\mathbf{u}$, so $\boldsymbol{\sigma}\mathbf{n} = \bar{\mathbf{t}}$ on s_σ, $\mathbf{u} = \bar{\mathbf{u}}$ on s_u and $\operatorname{div}\boldsymbol{\sigma} + \mathbf{b} = \rho\ddot{\mathbf{u}}$. Multiplying the equations of motion by $\mathbf{u}$ and integrating leads to

$$\int_v \rho\ddot{\mathbf{u}} \cdot \mathbf{u}dv = \int_v (\operatorname{div}\boldsymbol{\sigma} + \mathbf{b}) \cdot \mathbf{u}dv \tag{3.9.3}$$

Using the identity 1.14.16b, $\operatorname{div}(\mathbf{Av}) = \mathbf{v} \cdot \operatorname{div}\mathbf{A}^{\mathrm{T}} + \operatorname{tr}(\mathbf{A}\operatorname{grad}\mathbf{v})$, 1.10.10e, $\operatorname{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) = \mathbf{A} : \mathbf{B}$, and the symmetry of stress,

$$\int_v \rho\ddot{\mathbf{u}} \cdot \mathbf{u}dv = \int_v \{\operatorname{div}(\boldsymbol{\sigma}\mathbf{u}) - \boldsymbol{\sigma} : \operatorname{grad}(\mathbf{u}) + \mathbf{b} \cdot \mathbf{u}\}dv \tag{3.9.4}$$

and the divergence theorem 1.14.22c and Cauchy's law lead to

$$\int_s \mathbf{t} \cdot \mathbf{u}ds + \int_v \mathbf{b} \cdot \mathbf{u}dv = \int_v \boldsymbol{\sigma} : \operatorname{grad}\mathbf{u}dv + \int_v \rho\ddot{\mathbf{u}} \cdot \mathbf{u}dv \tag{3.9.5}$$

Splitting the surface integral into one over $s_{\mathbf{u}}$ and one over $s_{\boldsymbol{\sigma}}$ gives

$$\int_{s_{\mathbf{u}}} \mathbf{t} \cdot \bar{\mathbf{u}}ds + \int_{s_{\boldsymbol{\sigma}}} \bar{\mathbf{t}} \cdot \mathbf{u}ds + \int_v \mathbf{b} \cdot \mathbf{u}dv = \int_v \boldsymbol{\sigma} : \operatorname{grad}\mathbf{u}dv + \int_v \rho\ddot{\mathbf{u}} \cdot \mathbf{u}dv \tag{3.9.6}$$

Next, consider a second kinematically admissible displacement field $\mathbf{u}^*$, so $\mathbf{u}^* = \bar{\mathbf{u}}$ on s_u, which is completely arbitrary, in the sense that it is unrelated to either $\boldsymbol{\sigma}$ or $\mathbf{u}$. This time multiplying $\operatorname{div}\boldsymbol{\sigma} + \mathbf{b} = \rho\ddot{\mathbf{u}}$ across by $\mathbf{u}^*$, and following the same procedure, one arrives at

$$\int_{s_{\mathbf{u}}} \mathbf{t} \cdot \bar{\mathbf{u}}ds + \int_{s_{\boldsymbol{\sigma}}} \bar{\mathbf{t}} \cdot \mathbf{u}^*ds + \int_v \mathbf{b} \cdot \mathbf{u}^*dv = \int_v \boldsymbol{\sigma} : \operatorname{grad}\mathbf{u}^*dv + \int_v \rho\ddot{\mathbf{u}} \cdot \mathbf{u}^*dv \tag{3.9.7}$$

Let $\delta\mathbf{u} = \mathbf{u}^* - \mathbf{u}$, so the difference between $\mathbf{u}^*$ and $\mathbf{u}$ is infinitesimal, then subtracting 3.9.6 from 3.9.7 gives the principle of virtual work,

$$\boxed{\int_{s_{\boldsymbol{\sigma}}} \bar{\mathbf{t}} \cdot \delta\mathbf{u}ds + \int_v \mathbf{b} \cdot \delta\mathbf{u}dv = \int_v \boldsymbol{\sigma} : \operatorname{grad}(\delta\mathbf{u})dv + \int_v \rho\ddot{\mathbf{u}} \cdot \delta\mathbf{u}dv}$$

Principle of Virtual Work (spatial form) (3.9.8)

Note that since $\mathbf{u}, \mathbf{u}^*$, are kinematically admissible, $\delta\mathbf{u} = \mathbf{u}^* - \mathbf{u} = \mathbf{o}$ on s_u.

If one considers the $\mathbf{u}$ in 3.9.8 to be the actual displacement of the body, then $\delta\mathbf{u}$ can be considered to be a virtual displacement from the current configuration, Fig. 3.9.2. Again, it is emphasized that this virtual displacement leaves the stress, body force and applied traction unchanged.

One also has the transformed initial conditions: from 3.7.7d-e,

$$\int_v \mathbf{u}(\mathbf{x},\mathbf{t})_{t=0} \cdot \delta\mathbf{u}dv = \int_v \mathbf{u}_0(\mathbf{x}) \cdot \delta\mathbf{u}dv$$
$$\int_v \dot{\mathbf{u}}(\mathbf{x},\mathbf{t})_{t=0} \cdot \delta\mathbf{u}dv = \int_v \dot{\mathbf{u}}_0(\mathbf{x}) \cdot \delta\mathbf{u}dv \tag{3.9.9}$$

Eqns. 3.9.8 and 3.9.9 together constitute the **weak form** of the initial BVP 3.7.7.

The principle of virtual work can be grouped into three separate terms: the external virtual work:

$$\boxed{\delta W_{\text{ext}} = \int_{s_\sigma} \bar{\mathbf{t}} \cdot \delta\mathbf{u}ds + \int_v \mathbf{b} \cdot \delta\mathbf{u}dv} \quad \textbf{External Virtual Work} \tag{3.9.10}$$

the internal virtual work,

$$\boxed{\delta W_{\text{int}} = -\int_v \boldsymbol{\sigma} : \text{grad}(\delta\mathbf{u})dv} \quad \textbf{Internal Virtual Work} \tag{3.9.11}$$

and the virtual kinetic energy,

$$\boxed{\delta K = \int_v \rho\ddot{\mathbf{u}} \cdot \delta\mathbf{u}dv} \quad \textbf{Virtual Kinetic Energy} \tag{3.9.12}$$

corresponding to the statement 3.9.2.

Derivation from the Equations of Motion (Material Form)

The derivation in the spatial form follows exactly the same lines as for the spatial form.

This time, let $\mathbf{P}$ be a statically admissible stress field corresponding to a kinematically admissible displacement field $\mathbf{U}$, so $\mathbf{PN} = \overline{\mathbf{T}}$ on $S_\mathbf{P}$, $\mathbf{U} = \overline{\mathbf{U}}$ on $S_\mathbf{u}$ and $\text{div}\mathbf{P} + \mathbf{B} = \rho_0 \ddot{\mathbf{U}}$. This time one arrives at

$$\int_{S_\mathbf{P}} \overline{\mathbf{T}} \cdot \delta\mathbf{U}dS + \int_V \mathbf{B} \cdot \delta\mathbf{U}dV = \int_V \mathbf{P} : \text{Grad}(\delta\mathbf{U})dV + \int_V \rho_0 \ddot{\mathbf{U}} \cdot \delta\mathbf{U}dV \tag{3.9.13}$$

Again, one can consider $\mathbf{U}$ to be the actual displacement of the body, so that $\delta\mathbf{U}$ represents a virtual displacement from the current configuration. With

$$\mathbf{U} = \mathbf{x} - \mathbf{X}$$
$$\delta\mathbf{U} = \delta\mathbf{x} - \delta\mathbf{X} = \delta\mathbf{x} \tag{3.9.14}$$

the virtual work equation can be expressed in terms of the motion $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X})$,

$$\boxed{\int_{S_\mathbf{P}} \overline{\mathbf{T}} \cdot \delta\boldsymbol{\chi} dS + \int_V \mathbf{B} \cdot \delta\boldsymbol{\chi} dV = \int_V \mathbf{P} : \mathrm{Grad}(\delta\boldsymbol{\chi}) dV + \int_V \rho_0 \ddot{\mathbf{U}} \cdot \delta\boldsymbol{\chi} dV}$$

Principle of Virtual Work (material form) (3.9.15)

3.9.3 Principle of Virtual Work in terms of Strain Tensors

The principle of virtual work, in particular the internal virtual work term, can be expressed in terms of strain tensors.

Spatial Form

Using the commutative property of the variation 2.13.2, the term $\mathrm{grad}(\delta\mathbf{u})$ in the internal virtual work expression 3.9.8 can be written as

$$\begin{aligned} \mathrm{grad}(\delta\mathbf{u}) &= \frac{1}{2}\left(\mathrm{grad}(\delta\mathbf{u}) + \left(\mathrm{grad}(\delta\mathbf{u})\right)^{\mathrm{T}}\right) + \frac{1}{2}\left(\mathrm{grad}(\delta\mathbf{u}) - \left(\mathrm{grad}(\delta\mathbf{u})\right)^{\mathrm{T}}\right) \\ &= \delta\frac{1}{2}\left(\mathrm{grad}\mathbf{u} + \left(\mathrm{grad}\mathbf{u}\right)^{\mathrm{T}}\right) + \delta\frac{1}{2}\left(\mathrm{grad}\mathbf{u} - \left(\mathrm{grad}\mathbf{u}\right)^{\mathrm{T}}\right) \\ &= \delta\boldsymbol{\varepsilon} + \delta\boldsymbol{\Omega} \end{aligned} \qquad (3.9.16)$$

where $\boldsymbol{\varepsilon}$ is the (symmetric) small strain tensor and $\boldsymbol{\Omega}$ is the (skew-symmetric) small rotation tensor, Eqn 2.7.2. Using the fact that the double contraction of a symmetric tensor ($\boldsymbol{\sigma}$) and a skew-symmetric one ($\boldsymbol{\Omega}$) is zero, 1.10.31c, one has

$$\delta W_{\mathrm{int}} = -\int_v \boldsymbol{\sigma} : \mathrm{grad}(\delta\mathbf{u}) dv = -\int_v \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon}\, dv \qquad (3.9.17)$$

Thus the stresses do internal virtual work along the virtual strains $\delta\boldsymbol{\varepsilon}$. One has

$$\int_{s_\sigma} \bar{\mathbf{t}} \cdot \delta\mathbf{u} ds + \int_v \mathbf{b} \cdot \delta\mathbf{u} dv = \int_v \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon} dv + \int_v \rho\ddot{\mathbf{u}} \cdot \delta\mathbf{u} dv \qquad (3.9.18)$$

Note that, although the small strain has been introduced here, this formulation is not restricted to small-strain theory. It is only the virtual strains that must be infinitesimal – there is no restriction on the magnitude of the *actual* strains.

From 2.13.15, the Lie-variation of the Euler-Almansi strain $\mathbf{e}$ is $\delta_{\mathrm{L}}\mathbf{e} = \delta\boldsymbol{\varepsilon}$, so the internal; virtual work can be expressed as

$$\delta W_{\mathrm{int}} = -\int_v \boldsymbol{\sigma} : \delta_{\mathrm{L}}\mathbf{e}\, dv \qquad (3.9.19)$$

Material Form

From Eqn. 3.9.15 and Eqn. 2.13.9,

$$\delta W_{\text{int}} = -\int_V \mathbf{P} : \delta \mathbf{F} dV \tag{3.9.20}$$

so

$$\int_{S_\mathbf{P}} \overline{\mathbf{T}} \cdot \delta \boldsymbol{\chi} dS + \int_V \mathbf{B} \cdot \delta \boldsymbol{\chi} dV = \int_V \mathbf{P} : \delta \mathbf{F} dV + \int_V \rho_0 \ddot{\mathbf{U}} \cdot \delta \boldsymbol{\chi} dV \tag{3.9.21}$$

Derivation of the Material Form directly from the Spatial Form

To transform the spatial form of the virtual work equation into the material form, first note that, with 3.9.14b,

$$\boldsymbol{\sigma} : \text{grad}(\delta \mathbf{u}) = \boldsymbol{\sigma} : \text{grad}(\delta \mathbf{x}) \tag{3.9.22}$$

Then, using 2.2.8b, $\text{grad}\mathbf{v} = \text{Grad}\mathbf{V}\,\mathbf{F}^{-1}$, 2.13.9, $\delta \mathbf{F} = \text{Grad}(\delta \mathbf{u})$, 1.10.3h, $\mathbf{A} : (\mathbf{BC}) = \left(\mathbf{AC}^{\mathrm{T}}\right) : \mathbf{B}$, and 3.5.10, $\mathbf{P} = J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}$,

$$\begin{aligned} \boldsymbol{\sigma} : \text{grad}(\delta \mathbf{x}) &= \boldsymbol{\sigma} : \left(\text{Grad}(\delta \mathbf{x})\mathbf{F}^{-1}\right) \\ &= \boldsymbol{\sigma} : \left(\delta \mathbf{F}\mathbf{F}^{-1}\right) \\ &= \left(\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right) : \delta \mathbf{F} \\ &= \left(J^{-1}\mathbf{P}\right) : \delta \mathbf{F} \end{aligned} \tag{3.9.23}$$

which converts 3.9.17 into 3.9.120.

Also, again comparing 3.9.17 and 3.9.20, using the trace properties 1.10.10, and Eqns. 3.5.9 and 2.13.11b,

$$\mathbf{P} : \delta \mathbf{F} = J\boldsymbol{\sigma} : \delta \boldsymbol{\varepsilon} = J\text{tr}\left(\boldsymbol{\sigma}\delta\boldsymbol{\varepsilon}\right) = J\text{tr}\left(\mathbf{F}^{-1}\boldsymbol{\sigma}\delta\boldsymbol{\varepsilon}\mathbf{F}\right) = J\left(\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}\right) : \left(\mathbf{F}^{\mathrm{T}}\delta\boldsymbol{\varepsilon}\mathbf{F}\right) = \mathbf{S} : \delta \mathbf{E} \tag{3.9.24}$$

and so the internal work can also be expressed as an integral of $\mathbf{S} : \delta \mathbf{E}$ over the reference volume.

The Internal Virtual Work and Work Conjugate Tensors

The expressions for stress power 3.8.15, 3.8.29, and internal virtual work are very similar. For the material description, the time derivatives in the former are simply replaced with the variation to get the latter:

$$\mathbf{P} : \dot{\mathbf{F}} = \mathbf{S} : \dot{\mathbf{E}} \quad \rightarrow \quad \mathbf{P} : \delta \mathbf{F} = \mathbf{S} : \delta \mathbf{E} \tag{3.9.25}$$

For spatial tensors, the rate of strain tensor, e.g. **d**, is replaced with a Lie variation 2.13.14. For example, $J\boldsymbol{\sigma} : \mathbf{d} = J\boldsymbol{\sigma} : \mathrm{L}_{\mathrm{v}}^{b}\mathbf{e}$ (see 2.12.41-42) becomes:

$$J\boldsymbol{\sigma} : \mathbf{d} = J\boldsymbol{\sigma} : \mathrm{L}_{\mathrm{v}}^{b}\mathbf{e} \quad \rightarrow \quad J\boldsymbol{\sigma} : \delta_{\mathrm{L}}\mathbf{e} \tag{3.9.26}$$

3.9.4 Derivation of the Strong Form from the Weak Form

Just as the strong form (equations of motion and boundary conditions) was converted into the weak form (principle of virtual work), the weak form can be converted back into the strong form. For example,

$$\begin{aligned}\int_v \boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon} dv + \int_v \rho\ddot{\mathbf{u}} \cdot \delta\mathbf{u} dv &= \int_v \boldsymbol{\sigma} : \delta \text{grad}\mathbf{u} dv + \int_v \rho\ddot{\mathbf{u}} \cdot \delta\mathbf{u} dv \\ &= \int_v \left\{\text{div}\left(\boldsymbol{\sigma}\cdot\delta\mathbf{u}\right) - \left(\text{div}\boldsymbol{\sigma} - \rho\ddot{\mathbf{u}}\right)\cdot\delta\mathbf{u}\right\} dv \\ &= \int_s \mathbf{t}\cdot\delta\mathbf{u} ds - \int_v \left\{\left(\text{div}\boldsymbol{\sigma} - \rho\ddot{\mathbf{u}}\right)\cdot\delta\mathbf{u}\right\} dv \\ &= \int_{s_\sigma} \mathbf{t}\cdot\delta\mathbf{u} ds - \int_v \left\{\left(\text{div}\boldsymbol{\sigma} - \rho\ddot{\mathbf{u}}\right)\cdot\delta\mathbf{u}\right\} dv\end{aligned} \tag{3.9.27}$$

and the last line follows from the fact that $\delta\mathbf{u} = \mathbf{o}$ on $s_{\mathbf{u}}$. Thus the weak form now reads

$$\int_{s_\sigma} \left(\mathbf{t} - \bar{\mathbf{t}}\right)\cdot\delta\mathbf{u} ds - \int_v \left(\text{div}\boldsymbol{\sigma} + \mathbf{b} - \rho\ddot{\mathbf{u}}\right)\cdot\delta\mathbf{u} dv = 0 \tag{3.9.28}$$

and, since $\delta\mathbf{u}$ is arbitrary, one finds that the expressions in the parentheses are zero, and so 3.7.7 is recovered.

3.9.5 Conservative Systems

Thus far, no assumption has been made about the nature of the internal forces acting in the material. Indeed, the principle of virtual work applies to all types of materials.

Now, however, attention is restricted to the special case where the system is conservative, in the sense that the work done by the external loads and the internal forces can be written in terms of potential energy functions[1]. Further, for brevity, assume also that the material is in static equilibrium, i.e. the kinetic energy term is zero.

In other words, it is assumed that the internal virtual work term can be expressed in the form of a virtual potential energy function:

$$\int_v \boldsymbol{\sigma} : \text{grad}(\delta\mathbf{u}) dv = \int_v \delta U dv \tag{3.9.29}$$

Here, U is considered to be a function of $\mathbf{u}$, and the variation is to be understood as in Eqn. 2.13.5, $\delta U\left(\mathbf{u}, \delta\mathbf{u}\right) \equiv \partial_{\mathbf{u}} U[\delta\mathbf{u}]$.

If the loads can be regarded as functions of $\mathbf{u}$ only then, since they are conservative, they may be written as the gradient of a scalar potential:

[1] The external loads being conservative would exclude, for example, cases of frictional loading

$$\mathbf{b} = -\frac{\partial U_b}{\partial \mathbf{u}}, \quad \bar{\mathbf{t}} = -\frac{\partial U_t}{\partial \mathbf{u}} \tag{3.9.30}$$

Then, with

$$\delta U_b = \frac{\partial U_b}{\partial \mathbf{u}} \cdot \delta \mathbf{u}, \quad \delta U_t = \frac{\partial U_t}{\partial \mathbf{u}} \cdot \delta \mathbf{u} \tag{3.9.31}$$

and using the commutative property 2.13.3 of the variational operator, one arrives at

$$\delta \left\{ \int_v U dv + \int_{s_\sigma} U_t ds + \int_v U_b dv \right\} = \delta \overline{U}(\mathbf{u}) = 0 \tag{3.9.32}$$

The quantity inside the brackets is the total potential energy of the system. This statement is the **principle of stationary potential energy**: the value of the quantity inside the parentheses, i.e. $\overline{U}(\mathbf{u})$, is stationary at the true solution $\mathbf{u}$.

Eqn. 3.9.32 is an example of a **Variational Principle**, that is, a principle expressed in the form of a variation of a functional. Note that the principle of virtual work in the form 3.9.8 is not a variational principle, since it is not expressed as the variation of one functional.

Body Forces

Body forces can usually be expressed in the form 3.9.30. For example, with gravity loading, $\mathbf{b} = \rho \mathbf{g}$, where $\mathbf{g}$ is the constant acceleration due to gravity. Then $U_b = -\rho \mathbf{g} \cdot \mathbf{u}$ ($\delta \mathbf{b} = \delta \mathbf{g} = \mathbf{o}$ and $\mathbf{b} \cdot \delta \mathbf{u} = \delta(\mathbf{b} \cdot \mathbf{u})$, so $\int \mathbf{b} \cdot \delta \mathbf{u} dv = \delta \int \mathbf{b} \cdot \mathbf{u} dv$).

Material Form

In the material form, one again has a stationary principle if one can write $\mathbf{B} = -\partial U_B(\mathbf{U}) / \partial \mathbf{U}$, $\overline{\mathbf{T}} = -\partial U_T(\mathbf{U}) / \partial \mathbf{U}$ (or, equivalently, replacing $\mathbf{U}$ with the motion $\boldsymbol{\chi}$). In the case of dead loading, §3.7.1, $\overline{\mathbf{T}} = \overline{\mathbf{T}}(\mathbf{X})$ is independent of the motion so (similar to the case of gravity loading above) $U_T = -\overline{\mathbf{T}} \cdot \mathbf{u}$ with $\delta \overline{\mathbf{T}} = \mathbf{o}$ and $\int \overline{\mathbf{T}} \cdot \delta \boldsymbol{\chi} dV = \delta \int \overline{\mathbf{T}} \cdot \boldsymbol{\chi} dV$.

Deformation Dependent Traction

In many practical cases, the traction will depend on not only the motion, but also the strain. In that case, one can write

$$\int_{s_\sigma} \bar{\mathbf{t}} \cdot \delta \mathbf{u} ds = \int_s \bar{\boldsymbol{\sigma}} \mathbf{n} \cdot \delta \mathbf{u} ds = \int_s \mathbf{n} \bar{\boldsymbol{\sigma}} \delta \mathbf{u} ds = \int_v \operatorname{div}(\bar{\boldsymbol{\sigma}} \delta \mathbf{u}) dv = \int_v (\operatorname{div} \bar{\boldsymbol{\sigma}} \cdot \delta \mathbf{u} + \bar{\boldsymbol{\sigma}} : \operatorname{grad} \delta \mathbf{u}) dv$$

One might be able to then introduce a scalar function ϕ such that

$$\delta\phi(\mathbf{u},\boldsymbol{\varepsilon}) = \frac{\partial\phi}{\partial\mathbf{u}}\cdot\delta\mathbf{u} + \frac{\partial\phi}{\partial\boldsymbol{\varepsilon}}:\delta\boldsymbol{\varepsilon} \quad \text{with} \quad \frac{\partial\phi}{\partial\mathbf{u}} = \operatorname{div}\overline{\boldsymbol{\sigma}},\ \frac{\partial\phi}{\partial\boldsymbol{\varepsilon}} = \overline{\boldsymbol{\sigma}} \tag{3.9.33}$$

In the material form, one would have $\overline{\mathbf{T}} = \overline{\mathbf{P}}\mathbf{N}$ with

$$\int_{S_\mathbf{P}} \overline{\mathbf{T}}\cdot\delta\boldsymbol{\chi}dS = \int_V \left(\operatorname{Div}\overline{\mathbf{P}}\cdot\delta\boldsymbol{\chi} + \overline{\mathbf{P}}:\mathbf{F}\right)dV$$

and then one might be able to introduce a scalar function $\phi(\boldsymbol{\chi},\mathbf{F})$ such that

$$\delta\phi = \frac{\partial\phi}{\partial\boldsymbol{\chi}}\cdot\delta\boldsymbol{\chi} + \frac{\partial\phi}{\partial\mathbf{F}}:\delta\mathbf{F} \quad \text{with} \quad \frac{\partial\phi}{\partial\boldsymbol{\chi}} = \operatorname{Div}\overline{\mathbf{P}},\ \frac{\partial\phi}{\partial\mathbf{F}} = \overline{\mathbf{P}} \tag{3.9.34}$$

For example, considering again the fluid pressure example of §3.7.1, one can let $\phi = -pJ$ so that, using 1.15.7, $\partial\phi/\partial\mathbf{F} = -pJ\mathbf{F}^{-\mathrm{T}}$. Then $\overline{\mathbf{P}} = -p\mathbf{F}^{-\mathrm{T}} = \partial\phi/\partial\mathbf{F}\big|_{J=1}$, $\operatorname{Div}\overline{\mathbf{P}} = \rho g\mathbf{E}_2$ and $\int\overline{\mathbf{T}}\cdot\delta\boldsymbol{\chi}dS = -\delta\int pJdV$.

3.9.6 The Principle of Virtual Power

The principle of virtual power is similar to the principle of virtual work, the only difference between them being that a virtual velocity $\delta\mathbf{v}$ is used in the former rather than a virtual displacement. To derive the virtual power equation, multiply the equations of motion by the virtual velocity function, and integrate over the current configuration, giving

$$\begin{aligned}
\int_v \rho\frac{d\mathbf{v}}{dt}\cdot\delta\mathbf{v}dv = \int_v (\operatorname{div}\boldsymbol{\sigma} + \mathbf{b})\cdot\delta\mathbf{v}dv &= \int_v \left\{\operatorname{div}(\boldsymbol{\sigma}\delta\mathbf{v}) - \boldsymbol{\sigma}:\frac{\partial(\delta\mathbf{v})}{\partial\mathbf{x}} + \mathbf{b}\cdot\delta\mathbf{v}\right\}dv \\
&= \int_v \left\{\operatorname{div}(\boldsymbol{\sigma}\delta\mathbf{v}) - \boldsymbol{\sigma}:\delta\frac{\partial\mathbf{v}}{\partial\mathbf{x}} + \mathbf{b}\cdot\delta\mathbf{v}\right\}dv \\
&= \int_s \mathbf{t}\cdot\delta\mathbf{v}ds - \int_v \boldsymbol{\sigma}:\delta\mathbf{d}dv + \int_v \mathbf{b}\cdot\delta\mathbf{v}dv
\end{aligned} \tag{3.9.35}$$

These equations are identical to the mechanical balance equations 3.8.16, except that the actual velocity is replaced with a virtual velocity. The term $-\int_v \boldsymbol{\sigma}:\delta\mathbf{d}dv$ is called the **internal virtual power**.

Note that here, unlike the virtual displacement function in the work equation, the virtual velocity does not have to be infinitesimal. This can be seen more clearly if one derives this equation directly from the virtual work equation. If the infinitesimal virtual displacement $\delta\mathbf{u}$ occurs over an infinitesimal time interval δt, the virtual velocity is the finite quantity $\delta\mathbf{u}/\delta t$, which here is labelled $\delta\mathbf{v}$. The virtual power equation can thus be obtained by dividing the virtual work equation through by δt.

Again, supposing that the velocities are specified over that part of the surface s_v and tractions over $s_{\boldsymbol{\sigma}}$, the principle of virtual power can be written for the case of a kinematically admissible virtual velocity field:

$$\boxed{\int_{s_{\boldsymbol{\sigma}}} \bar{\mathbf{t}} \cdot \delta\mathbf{v} ds + \int_v \mathbf{b} \cdot \delta\mathbf{v} dv = \int_v \boldsymbol{\sigma} : \delta\mathbf{d} dv + \int_v \rho \frac{d\mathbf{v}}{dt} \cdot \delta\mathbf{v} dv} \quad \textbf{Principle of Virtual Power} \qquad (3.9.36)$$

In words, the principle of virtual power states that *at any time t, the total virtual power of the external, internal and inertia forces is zero in any admissible virtual state of motion*.

3.9.7 Linearisation of the Internal Virtual Work

In order to solve the virtual work equations in anything but the most simple cases, one must apply some approximate numerical method. This will usually involve linearising the non-linear virtual work equations. To this end, the internal virtual work term will be linearised in what follows.

Material Description

In the material description, one has

$$\delta W_{\text{int}} = \int_V \mathbf{S}(\mathbf{E}(\mathbf{u})) : \delta\mathbf{E}(\mathbf{u}) dV \qquad (3.9.37)$$

in which the Green-Lagrange strain is considered to be a function of the displacement, Eqn. 2.2.46, and the PK2 stress is a function of the Green-Lagrange strain; the precise functional dependence of **S** on **E** will depend on the material under study (see Part IV).

The linearisation of the variation of a function is given by (see §2.13.2)

$$\mathrm{L}\ \delta W_{\text{int}}(\mathbf{u}, \Delta\mathbf{u}) = \delta W_{\text{int}}(\mathbf{u}) + \Delta\delta W_{\text{int}}(\mathbf{u}, \Delta\mathbf{u}) \qquad (3.9.38)$$

where

$$\begin{aligned}
\Delta\delta W_{\text{int}}(\mathbf{u}, \Delta\mathbf{u}) &= \partial_{\mathbf{u}} \delta W_{\text{int}}[\Delta\mathbf{u}] \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \delta W_{\text{int}}(\mathbf{u} + \varepsilon\Delta\mathbf{u}) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_V \mathbf{S}(\mathbf{E}(\mathbf{u} + \varepsilon\Delta\mathbf{u})) : \delta\mathbf{E}(\mathbf{u} + \varepsilon\Delta\mathbf{u}) dV \\
&= \int_V \left\{ \mathbf{S}(\mathbf{E}(\mathbf{u})) : \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \delta\mathbf{E}(\mathbf{u} + \varepsilon\Delta\mathbf{u}) + \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{S}(\mathbf{E}(\mathbf{u} + \varepsilon\Delta\mathbf{u})) : \delta\mathbf{E}(\mathbf{u}) \right\} dV \\
&= \int_V \{ \mathbf{S}(\mathbf{E}(\mathbf{u})) : \Delta\delta\mathbf{E}(\mathbf{u}, \Delta\mathbf{u}) + \Delta\mathbf{S}(\mathbf{E}(\mathbf{u}, \Delta\mathbf{u})) : \delta\mathbf{E}(\mathbf{u}) \} dV
\end{aligned}$$

(3.9.39)

The linearization of the variation of the Green-Lagrange strain is given by 2.13.24, $\Delta\delta\mathbf{E} = \mathrm{sym}\left((\mathrm{Grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right)$. With the PK2 stress symmetric, one has, with 1.10.3h, 1.10.31c,

$$\begin{aligned}\mathbf{S}(\mathbf{E}(\mathbf{u})):\Delta\delta\mathbf{E}(\mathbf{u},\Delta\mathbf{u}) &= \mathbf{S}(\mathbf{E}(\mathbf{u})):\mathrm{sym}\left((\mathrm{Grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right)\\ &= \mathbf{S}(\mathbf{E}(\mathbf{u})):(\mathrm{Grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\\ &= \mathrm{Grad}\delta\mathbf{u}:(\mathrm{Grad}\Delta\mathbf{u})\mathbf{S}(\mathbf{E}(\mathbf{u}))\end{aligned} \tag{3.9.40}$$

For the second term in 3.9.39, from 2.13.22, the variation of $\mathbf{E}$ is

$$\begin{aligned}\delta\mathbf{E} &= \tfrac{1}{2}\left[(\mathrm{Grad}\delta\mathbf{u})^{\mathrm{T}}\mathbf{F} + \mathbf{F}^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right]\\ &= \tfrac{1}{2}\left[(\mathbf{F}^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u})^{\mathrm{T}} + \mathbf{F}^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right]\\ &= \mathrm{sym}\left(\mathbf{F}^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}\right)\end{aligned} \tag{3.9.41}$$

What remains is the calculation of the linearisation of the PK2 stress. One has using the chain rule,

$$\begin{aligned}\Delta\mathbf{S}(\mathbf{E}(\mathbf{u},\Delta\mathbf{u})) &= \partial_{\mathbf{u}}\mathbf{S}[\Delta\mathbf{u}]\\ &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\mathbf{S}(\mathbf{E}(\mathbf{u}+\varepsilon\Delta\mathbf{u}))\\ &= \frac{\partial\mathbf{S}(\mathbf{E})}{\partial\mathbf{E}}:\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\mathbf{E}((\mathbf{u}+\varepsilon\Delta\mathbf{u}))\\ &= \frac{\partial\mathbf{S}(\mathbf{E})}{\partial\mathbf{E}}:\Delta\mathbf{E}(\mathbf{u},\Delta\mathbf{u})\end{aligned} \tag{3.9.42}$$

Denote the fourth order tensor $\partial\mathbf{S}(\mathbf{E})/\partial\mathbf{E}$ by $\mathbf{C}$ and assume that it has the minor symmetries1.12.10. Then (see 3.9.41), with

$$\Delta\mathbf{E} = \mathrm{sym}\left(\mathbf{F}^{\mathrm{T}}\mathrm{Grad}\Delta\mathbf{u}\right) \tag{3.9.43}$$

the linear increments in 3.9.38 become

$$\boxed{\begin{aligned}\Delta\delta W_{\mathrm{int}}(\mathbf{u},\Delta\mathbf{u}) &= \int_V \{\mathrm{Grad}\delta\mathbf{u}:(\mathrm{Grad}\Delta\mathbf{u})\mathbf{S}(\mathbf{E}(\mathbf{u}))\\ &\qquad\qquad + \mathbf{F}^{\mathrm{T}}\mathrm{Grad}\delta\mathbf{u}:\mathbf{C}:\mathbf{F}^{\mathrm{T}}\mathrm{Grad}\Delta\mathbf{u}\}dV\\ \Delta\delta W_{\mathrm{int}}(\mathbf{u},\Delta\mathbf{u}) &= \int_V \left\{\frac{\partial\delta u_i}{\partial X_b}\frac{\partial\Delta u_i}{\partial X_d}S_{bd} + F_{ka}\frac{\partial\delta u_k}{\partial X_b}C_{abcd}F_{jc}\frac{\partial\Delta u_j}{\partial X_d}\right\}dV\\ &= \int_V \frac{\partial\delta u_i}{\partial X_b}\{\delta_{ij}S_{bd} + F_{ia}F_{jc}C_{abcd}\}\frac{\partial\Delta u_j}{\partial X_d}dV\end{aligned}} \tag{3.9.44}$$

The first term is due to the current stress and is called the (**initial**) **stress contribution**. The second term depends on the material properties and is called the **material contribution**. Solution formulations based on 3.9.44 are called **total Lagrangian**.

Spatial Description

The spatial description can be obtained by pushing forward the material description. First note that the linearization of the Kirchhoff stress is, from 3.5.13,

$$\begin{aligned} \mathrm{L}\ \boldsymbol{\tau}(\mathbf{u},\Delta\mathbf{u}) &= \chi_*\left(\mathrm{L}\ \mathbf{S}(\mathbf{u},\Delta\mathbf{u})\right)^{\#} \\ &= \chi_*\left(\mathbf{S}(\mathbf{u},\Delta\mathbf{u})\right)^{\#} + \chi_*\left(\Delta\ \mathbf{S}(\mathbf{u},\Delta\mathbf{u})\right)^{\#} \\ &= \boldsymbol{\tau}(\mathbf{u}) + \mathbf{F}\Delta\mathbf{S}(\mathbf{u},\Delta\mathbf{u})\mathbf{F}^{\mathrm{T}} \end{aligned} \tag{3.9.45}$$

so that, as in the derivation of the material term in 3.9.44, and using 2.4.8,

$$\begin{aligned} \Delta\boldsymbol{\tau}(\mathbf{u},\Delta\mathbf{u}) &= \mathbf{F}\left(\mathbf{C} : \mathbf{F}^{\mathrm{T}}\mathrm{grad}\Delta\mathbf{u}\mathbf{F}\right)\mathbf{F}^{\mathrm{T}} \\ \Delta\boldsymbol{\tau}(\mathbf{u},\Delta\mathbf{u}) &= F_{ia}F_{jb}F_{kc}F_{ld}C_{abcd}\frac{\partial \Delta u_k}{\partial x_l} \end{aligned} \tag{3.9.46}$$

Define the fourth-order spatial tensor $\mathbf{c}$ through

$$c_{ijkl} = J^{-1}F_{ia}F_{jb}F_{kc}F_{ld}C_{abcd} \tag{3.9.47}$$

so that

$$\Delta\boldsymbol{\tau}(\mathbf{u},\Delta\mathbf{u}) = J\,\mathbf{c} : \mathrm{grad}\Delta\mathbf{u} \tag{3.9.48}$$

Then, from 3.9.39,

$$\boxed{\begin{aligned} \Delta\delta W_{\mathrm{int}}(\mathbf{u},\Delta\mathbf{u}) &= \int_V \left\{\chi_*(\mathbf{S})^{\#} : \chi_*(\Delta\delta\mathbf{E})^{b} + \chi_*(\Delta\mathbf{S})^{\#} : \chi_*(\delta\mathbf{E})^{b}\right\} dV \\ &= \int_V \left\{\boldsymbol{\tau} : \mathrm{sym}\left((\mathrm{grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{grad}\delta\mathbf{u}\right) + J\,\mathbf{c} : \mathrm{grad}\Delta\mathbf{u} : \mathrm{sym}(\mathrm{grad}\delta\mathbf{u})\right\} dV \\ &= \int_v \left\{\boldsymbol{\sigma} : (\mathrm{grad}\Delta\mathbf{u})^{\mathrm{T}}\mathrm{grad}\delta\mathbf{u} + \mathbf{c} : \mathrm{grad}\Delta\mathbf{u} : \mathrm{grad}\delta\mathbf{u}\right\} dv \\ &= \int_v \left\{\mathrm{grad}\delta\mathbf{u} : \mathrm{grad}\Delta\mathbf{u}\,\boldsymbol{\sigma} + \mathrm{grad}\delta\mathbf{u} : \mathbf{c} : \mathrm{grad}\Delta\mathbf{u}\right\} dv \\ \\ \Delta\delta W_{\mathrm{int}}(\mathbf{u},\Delta\mathbf{u}) &= \int_v \left\{\frac{\partial \delta u_a}{\partial x_b}\frac{\partial \Delta u_a}{\partial x_d}\sigma_{bd} + \frac{\partial \delta u_a}{\partial x_b} c_{abcd}\frac{\partial \Delta u_c}{\partial x_d}\right\} dv \\ &= \int_v \frac{\partial \delta u_a}{\partial x_b}\left\{\delta_{ac}\sigma_{bd} + c_{abcd}\right\}\frac{\partial \Delta u_c}{\partial x_d}\, dv \end{aligned}} \tag{3.9.49}$$

Solution formulations based on 3.9.49 are called **updated-Lagrangian**.

3.10 Convected Coordinates

Some of the important results from sections 3.1-3.9 are now re-expressed in terms of convected coordinates. As before, any relations expressed in symbolic form hold also in the convected coordinate system.

3.10.1 The Stress Tensors

Traction and Stress Components

Consider a differential parallelepiped element in the current configuration bounded by the coordinate curves as in Fig. 3.10.1 (see Fig. 1.16.2). The bounding vectors are $d\Theta^1\mathbf{g}_1, d\Theta^2\mathbf{g}_2$ and $d\Theta^3\mathbf{g}_3$. The surface area $d\bar{S}_1$ of a face of the elemental parallelepiped on which Θ_1 is constant, to which $\mathbf{g}^1$ is normal, is then given by Eqn. 1.16.35,

$$d\bar{S}_1 = \sqrt{g\,g^{11}}\,d\Theta^2 d\Theta^3 \tag{3.10.1}$$

and similarly for the other surfaces.

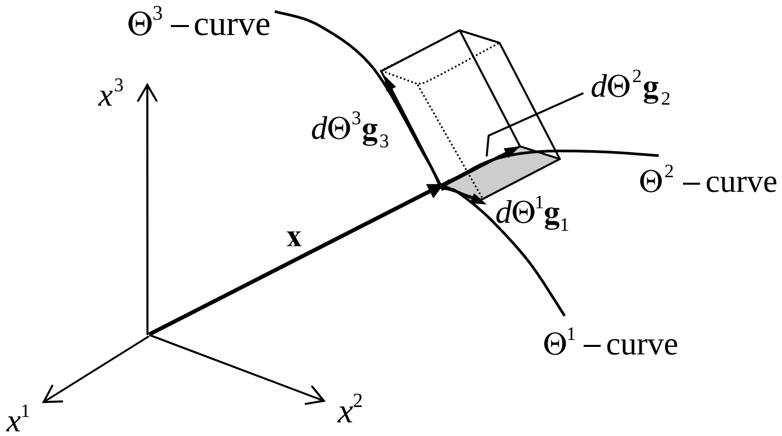


Figure 3.10.1: vector elements bounding surface elements

The **positive side** of a face is defined as that whose outward normal is in the direction of the associated contravariant base vector. The unit normal $\bar{\mathbf{n}}^i$ to a positive side is the same as the unit contravariant base vector; as in Eqn. 1.16.14,

$$\bar{\mathbf{n}}^i = \hat{\mathbf{g}}^i = \frac{\mathbf{g}^i}{\sqrt{g^{ii}}} \qquad (\text{no sum}) \tag{3.10.2}$$

Let the force $d\mathbf{F}^i$ acting on the surface element with normal $\bar{\mathbf{n}}^i$ be $d\bar{S}_i\mathbf{t}^{\langle i\rangle}$ (no sum over i), Fig. 3.10.2, so that $\mathbf{t}^{\langle i\rangle}$ is the traction (force per unit area).

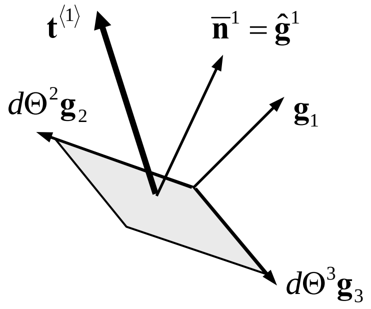


Figure 3.10.2: traction acting on a surface element

The components of $\mathbf{t}^{\langle i\rangle}$ along the unit covariant base vectors are denoted by $\sigma^{\langle ji\rangle}$:

$$\mathbf{t}^{\langle i\rangle} = \sigma^{\langle ji\rangle}\hat{\mathbf{g}}_j = \sigma^{\langle ji\rangle}\frac{1}{\sqrt{g_{jj}}}\mathbf{g}_j \qquad (3.10.3)$$

with no sum over the j in the $\sqrt{g_{jj}}$ term; $\sigma^{\langle ji\rangle}$ are called the **physical stress components**, Fig. 3.10.3.

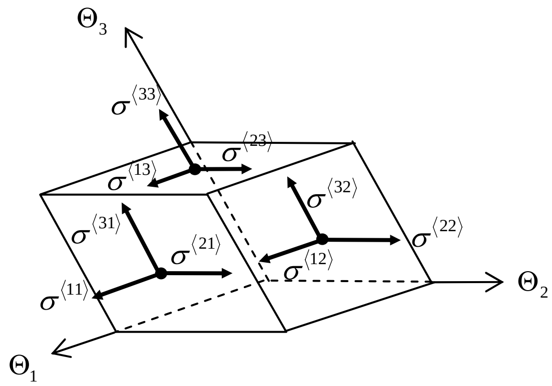


Figure 3.10.3: physical stress components

Introduce now a new vector $\mathbf{t}^i$ defined by

$$\mathbf{t}^i = \sqrt{g^{ii}}\,\mathbf{t}^{\langle i\rangle} \qquad \text{(no sum over } i\text{)} \qquad (3.10.4)$$

It will be shown that this vector is contravariant, that is, transforms between coordinate systems according to 1.17.3a (and so $\mathbf{t}^{\langle i\rangle}$ does not satisfy the vector transformation rule, hence the superscript in pointed brackets). The components of $\mathbf{t}^i$ along the covariant base vectors are denoted by σ^{ji}:

$$\mathbf{t}^i = \sigma^{ji}\mathbf{g}_j \qquad (3.10.5)$$

Comparing 3.10.3-5,

$$\sigma^{\langle ji \rangle} = \sqrt{\frac{g_{jj}}{g^{ii}}} \sigma^{ji} \qquad \text{(no sum)} \tag{3.10.6}$$

Cauchy's Law and the Cauchy Stress Tensor

Cauchy's law can now be derived in the same way as in §3.3, by considering a small tetrahedral free-body, Fig. 3.10.4. The physical stress components $\sigma^{\langle ij \rangle}$ shown act on the negative sides of the surfaces and so act in directions opposite that of the corresponding components on the positive sides (a consequence of Cauchy's Lemma). It is required to determine the traction **t** in terms of the physical stress components and the unit normal **n** to the base area.

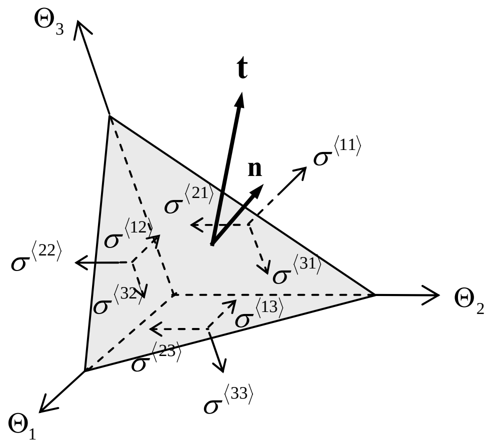


Figure 3.10.4: free body diagram of a tetrahedral portion of material

The normal to the base has components

$$\mathbf{n} = n^i \mathbf{g}_i = n_i \mathbf{g}^i \tag{3.10.7}$$

Consider the vector elements $d\mathbf{a}$ and $d\mathbf{b}$ shown in Fig. 3.10.5. Define the surface area element $d\mathbf{S}$ to be the vector with magnitude equal to twice the area of the tetrahedron base and in the direction of the normal to the base, so

$$\begin{aligned} \tfrac{1}{2} d\mathbf{S} = \tfrac{1}{2} dS\,\mathbf{n} &= \tfrac{1}{2}\left(d\mathbf{a} \times d\mathbf{b}\right) \\ &= \tfrac{1}{2}\left(\left(d\Theta^1 \mathbf{g}_1 - d\Theta^3 \mathbf{g}_3\right) \times \left(d\Theta^2 \mathbf{g}_2 - d\Theta^3 \mathbf{g}_3\right)\right) \\ &= \tfrac{1}{2}\left(d\Theta^1 d\Theta^2 \mathbf{g}_1 \times \mathbf{g}_2 + d\Theta^2 d\Theta^3 \mathbf{g}_2 \times \mathbf{g}_3 + d\Theta^3 d\Theta^1 \mathbf{g}_3 \times \mathbf{g}_1\right) \\ &= \tfrac{1}{2}\left(d\overline{\mathbf{S}}_1 + d\overline{\mathbf{S}}_2 + d\overline{\mathbf{S}}_3\right) \end{aligned} \tag{3.10.8}$$

where $d\overline{\mathbf{S}}_1$, $d\overline{\mathbf{S}}_2$, $d\overline{\mathbf{S}}_3$ are the surface element areas of the three coordinate sides of the parallelepiped of Fig. 3.10.1 (twice the area of the coordinate sides of the tetrahedron); from 3.10.2,

$$
\begin{aligned}
d\mathbf{S} &= d\overline{\mathbf{S}}_1 + d\overline{\mathbf{S}}_2 + d\overline{\mathbf{S}}_3 \\
dS\ \mathbf{n} &= d\overline{S}_i\ \overline{\mathbf{n}}^i \\
&= d\overline{S}_i \frac{1}{\sqrt{g^{ii}}} \mathbf{g}^i
\end{aligned} \tag{3.10.9}
$$

with no sum over the i in the $\sqrt{g^{ii}}$ term, or

$$
dS\ n_i \sqrt{g^{ii}} \mathbf{g}^i = d\overline{S}_i \mathbf{g}^i \tag{3.10.10}
$$

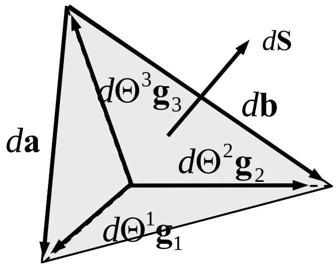


Figure 3.10.5: vector element of area for the base of the tetrahedron

The principle of linear momentum, in vector form, is then (cancelling out a factor of ½)

$$
\mathbf{t}\ dS - \mathbf{t}^{\langle i \rangle} d\overline{S}_i = 0 \tag{3.10.11}
$$

From 3.10.4,

$$
\mathbf{t}\ dS = \mathbf{t}^i \frac{1}{\sqrt{g^{ii}}} d\overline{S}_i = \mathbf{t}^i dS n_i \tag{3.10.12}
$$

and so

$$
\mathbf{t} = \mathbf{t}^i n_i = \sigma^{ji} n_i \mathbf{g}_j \tag{3.10.13}
$$

Defining the (symmetric) Cauchy stress tensor $\boldsymbol{\sigma}$ through

$$
\boxed{\boldsymbol{\sigma} = \sigma^{ij} \mathbf{g}_i \otimes \mathbf{g}_j} \qquad \textbf{Cauchy Stress Tensor} \tag{3.10.14}
$$

one arrives at Cauchy's law $\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}$.

The Cauchy stress is naturally a contravariant tensor because the normal vector upon which it operates to produce the traction is naturally represented in the form of a covariant vector (see 3.10.2).

Note that the stress can also be expressed in the form

$$\boldsymbol{\sigma} = \mathbf{t}^i \otimes \mathbf{g}_j \tag{3.10.15}$$

Other Stress Tensors

The PK1, PK2 and Kirchhoff stress tensors are

$$\begin{aligned} \mathbf{P} &= P^{ij}\mathbf{G}_i \otimes \mathbf{G}_j \\ \mathbf{S} &= S^{ij}\mathbf{G}_i \otimes \mathbf{G}_j \\ \boldsymbol{\tau} &= \tau^{ij}\mathbf{g}_i \otimes \mathbf{g}_j \end{aligned} \tag{3.10.16}$$

By definition, $\boldsymbol{\tau} = J\boldsymbol{\sigma}$, and so $\tau^{ij} = J\sigma^{ij}$. By definition, $\mathbf{S} = J\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}$, and so, from 2.9.8,

$$S^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = J\sigma^{ij}\mathbf{F}^{-1}\mathbf{g}_i \otimes \mathbf{F}^{-1}\mathbf{g}_j = J\sigma^{ij}\mathbf{G}_i \otimes \mathbf{G}_j = \tau^{ij}\mathbf{G}_i \otimes \mathbf{G}_j \tag{3.10.17}$$

Thus, as seen already, the Kirchhoff stress is the push-forward of the PK2 stress.

Similarly, by definition $\mathbf{P} = J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}$ and so

$$\begin{aligned} P^{ij}\mathbf{G}_i \otimes \mathbf{G}_j &= J\sigma^{kj}\mathbf{g}_k \otimes \mathbf{g}_j\mathbf{F}^{-1} \\ &= J\sigma^{kj}\mathbf{g}_k \otimes \mathbf{G}_j \\ &= J\sigma^{kj}\mathbf{F}\mathbf{G}_k \otimes \mathbf{G}_j \\ &= J\sigma^{kj}\left(F^{i}_{\cdot m}\mathbf{G}_i \otimes \mathbf{G}^m\right)\mathbf{G}_k \otimes \mathbf{G}_j \\ &= J\sigma^{kj}F^{i}_{\cdot k}\mathbf{G}_i \otimes \mathbf{G}_j \end{aligned} \tag{3.10.18}$$

3.10.2 The Equations of Motion

The Equations of motion have been given in the symbolic form by 3.6.2 and 3.6.9. To express these in curvilinear coordinates, recall the definition of the divergence of a tensor, 1.18.28,

$$\operatorname{div}\boldsymbol{\sigma} = \frac{\partial\boldsymbol{\sigma}}{\partial\Theta^k}\mathbf{g}^k = \sigma^{ij}|_k \left(\mathbf{g}_i \otimes \mathbf{g}_j\right)\mathbf{g}^k = \sigma^{ij}|_j\ \mathbf{g}_i \tag{3.10.19}$$

The spatial and material descriptions of the equations of motion are then

$$\boxed{\begin{aligned} \sigma^{ij}|_j\ \mathbf{g}_i + b^i\mathbf{g}_i &= \rho\frac{d\left(v^i\mathbf{g}_i\right)}{dt} \\ P^{ij}|_j\ \mathbf{G}_i + B^i\mathbf{G}_i &= \rho_0\frac{dV^i}{dt}\mathbf{G}_i \end{aligned}} \quad \textbf{Equations of Motion} \tag{3.10.20}$$

4 Fundamentals of Continuum Thermomechanics

In this Chapter, the laws of thermodynamics are reviewed and formulated for a continuum. The classical theory of thermodynamics, which is concerned with simple compressible systems, is discussed in sections 4.1-4.3, wherein are discussed the concepts of entropy, entropy production and entropy supply, the second law, the notions of reversibility and irreversibility, the thermodynamic potential functions (internal energy, enthalpy and the Gibbs and Helmholtz free energies). Continuum thermomechanics is discussed in section 4.4.

4.1 Classical Thermodynamics: The First Law

As an introduction to the thermomechanics of continua, in this section particularly simple materials undergoing simple deformation and/or heat-transfer processes are considered.

4.1.1 Properties and States

First, here is some essential terminology used to describe thermodynamic processes.

A **property** of a substance is a macroscopic characteristic to which a numerical value can be assigned at a given time. Thus, for example, the mass, volume and energy of a material, or the stress acting on a material, are properties. Work, on the other hand, is not a property, since a material does not "have a certain amount of work" (see the section which follows).

The **state** of a material is the condition of the material as described by its properties. For example a material which has properties volume V_1 and temperature θ_1 could be said to be in state '1' whereas if at some later time it has different properties V_2 and θ_2, it could be said to be in a different state, state '2'.

4.1.2 Work and Path Dependence

For the present purposes, a **system** can be defined to be a certain amount of matter which has fixed or movable boundaries. The state of a system can then be defined by assigning to it properties such as volume, pressure and so on. As will be seen, there are then two ways in which the state of the system can be changed, by interactions with its surroundings through **heat** or through **work**. The notion of heat, although familiar to us, will be defined precisely when the first law of thermodynamics is introduced below. First, consider the system shown below in Fig. 4.1.1, which consists of a block attached to an elastic spring, sliding over a rough surface. A force is applied to the "system" (denoted by the dotted line).

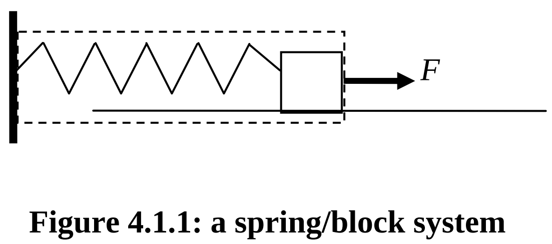


Figure 4.1.1: a spring/block system

The current state of the system can be described by the property *x*, the extension of the spring from its equilibrium position (and its velocity). However, the work done in moving the system from a previous state to the current state is unknown, since the block may have moved directly from its initial position to its current position, or it may have moved over and back many times before reaching the current position. Therefore the work done is **path dependent**. There are many different amounts of work which can be carried out to move a system from one state to another.

4.1.3 Thermal Equilibrium and Adiabatic Processes

Before getting to the first law, it is helpful to consider the notions of **thermal equilibrium** and **adiabatic** processes.

Thermal Equilibrium

Consider the following experiment: two blocks of copper, one of which our senses tell us is "warmer" than the other, are brought into contact and isolated from their surroundings, Fig. 4.2.1a. A number of observations would be made, for example:

(1) the volume of the warmer body decreases with time whereas the volume of the colder body increases, until no further changes take place and the bodies feel equally warm

(2) the electrical resistance of the warmer block decreases with time whereas that of the colder block increases, until the electrical resistances would become constant also.

When these and all such changes in observable properties cease, the interaction is at an end. One says that the two blocks are then in *thermal equilibrium*. In everyday language, one would say that the two blocks have the same **temperature**[1].

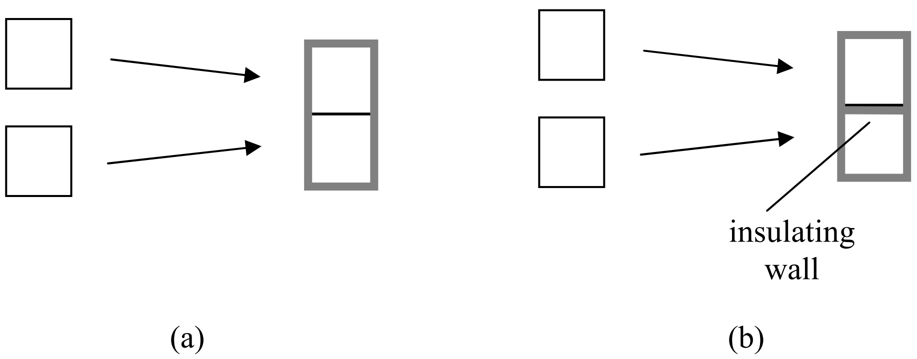


Figure 4.2.1: two blocks of copper brought into contact; (a) no insulating wall, (b) insulating wall

Adiabatic Conditions

Suppose now that, before the blocks are brought together, an **insulating wall** is put in place to separate them, Fig. 4.2.1b. By this is meant that the volume, electrical resistance, etc. of one block does not affect those of the other block. Again, in everyday language, one would simply say that the temperature of one block does not affect the temperature of the other. The term **adiabatic** is used to describe this situation.

[1] formally, temperature is defined through the **zeroth law of thermodynamics**, which states that if two systems are separately in thermal equilibrium with a third system, then they must be in thermal equilibrium with one another. This statement is tacitly assumed in every measurement of temperature – the third system being the thermometer

4.1.4 The First Law of Thermodynamics

The Experiments of Joule

James Joule carried out some ingenious experiments into the nature of work and heat transfer in materials in the 1840s. In his most famous experiment, Joule filled a container with a fluid and used a rotating paddle wheel, driven by falling weights, to stir the water. The container was thermally insulated and so the process was adiabatic. Joule measured the consequent rise in temperature of the fluid and noted that this change in the fluid's properties was due to the work done by the falling weights.

Further experiments were carried out in which raising the temperature of the thermally insulated fluid was induced by carrying out the necessary work in different ways, for example using electrical means. In all cases, the work required to raise the temperature by a fixed amount was the same.

The series of experiments showed that if a material is thermally insulated, there is only *one* amount of work which brings the material from one state to a second state. If one knows the first state and the second state, one knows the amount of work required to effect the change in state – the work is *path independent*.

It took many years for investigators to absorb the meaning of this experimental result; it was eventually accepted that there must exist a function U, a property of the system, such that

$$W = \Delta U = U_2 - U_1 \qquad \text{(adiabatic process)} \tag{4.1.1}$$

U is the **internal energy**, and the difference in internal energy between state 2 and state 1 is *defined* as equal to the work done in going from 1 to 2 by *adiabatic* means.

In the case of the stirred fluid, the increase in internal energy is due to the more rapidly moving fluid particles, that is, is equivalent to the increase in kinetic energy of the fluid particles.

Consider again the elastic spring system of Fig. 4.1.1, but now without the sliding over a rough surface, and completely thermally insulated. The work done now depends only on the current state (it equals $\frac{1}{2}kx^2 - \frac{1}{2}kx_0^2$, where k is the spring constant and x_0 is the equilibrium position). The internal energy of the system is seen to be in this case equivalent to the elastic potential energy in the spring.

The First Law

One can imagine now a careful experiment in which a material is thermally insulated from its surroundings and deformed through the work of a set of forces. The material can be deformed into different states, Fig. 4.1.3. The internal energy U will in general be different in each state. U could be measured by carefully recording the work done on the material to reach a given state.

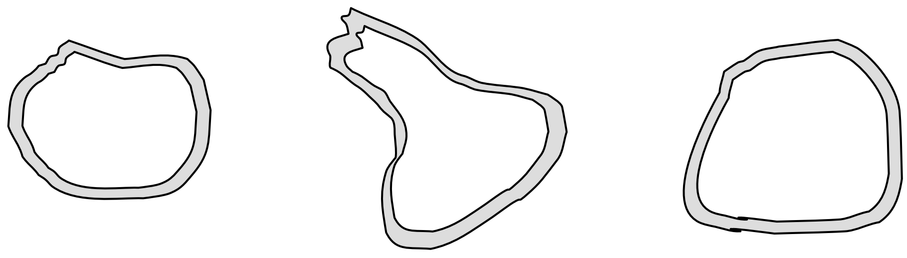


Figure 4.1.3: a thermally insulated material in three different states

Suppose that the internal energy of a material is known at various different states, through the conduction of the aforementioned experiment, in particular one knows the internal energy for the material at two given states, 1 and 2. Relax now the condition that the changes are adiabatic. What this means is that if one now brings the material into contact with another body, the properties of the material *can* be affected. Work is again done to take the material from state 1 to state 2 but it will now be found that, in general,

$$W \neq \Delta U = U_2 - U_1 \tag{4.1.2}$$

The difference between ΔU and W is *defined* as a measure of the *heat Q* which has entered the system in the change. Thus

$$\boxed{W + Q = \Delta U} \qquad \textbf{First Law of Thermodynamics} \tag{4.1.3}$$

This is **the first law of thermodynamics**. In words, t*he change in the internal energy is the sum of the work done plus the heat supplied*.

Note that the concept of heat Q (and internal energy) is introduced and defined with the first law. Like work, heat is a form of energy *transfer*; a body does not *contain* heat. Work is any means of changing the energy of a system other than heat.

Sign Convention for Work and Energy

The following sign convention will be used[2]

$$\begin{array}{ll} Q > 0 \;- & \text{heat enters the system} \\ Q < 0 \;- & \text{heat leaves the system} \\ W > 0 \;- & \text{work done on the system} \\ W < 0 \;- & \text{work done by the system} \end{array} \tag{4.1.4}$$

[2] many authors use the exact opposite sign convention for work as used here

Other types of Energy

When there are other energies involved, the first law must be amended. For a material moving with a certain velocity, one must also consider its kinetic energy, and the first law reads

$$W + Q = \Delta U + \Delta K \tag{4.1.5}$$

Other types of energy can be incorporated, for example gravitational potential energy and chemical energy[3]. All the different types of energy are often denoted simply by E, so the first law in general reads $W + Q = \Delta E$.

Inside the Black Box

In this continuum treatment of thermodynamics (or **phenomenological thermodynamics**), it is not necessary to look inside and consider the billions of molecules inside the "black box" of a system. However, it is helpful to think of the molecules of a material as having certain micro-velocities and it is the mean velocity of these micro-velocities which manifests itself as the macroscopic velocity property, and the statistical fluctuations of the micro-velocities from the mean velocity are assumed to cancel out, Fig. 4.1.4.

The micro-velocity fluctuations give rise to an *internal* kinetic energy which manifests itself as the macroscopic temperature, as in the stirred fluid mentioned above. The interaction between the elementary particles and the surroundings of the element causes energy to be transferred to the surroundings. This is the **heat flow** through the boundary of the system. This energy exchange can occur even when the shape of the element does not change, whereas a change in potential energy implies a deformation which will induce a re-arrangement of the molecules and change in shape or volume of the system.

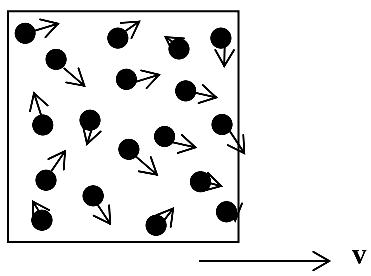


Figure 4.1.4: a system moving with velocity v

The property of pressure or stress of the system is by definition determined by the forces exerted by the elementary particles around the boundary. The fluctuations and micro-movement of the elementary particles will cause stress fluctuations but again these are assumed to cancel out.

[3] a potential energy which can be accessed when molecular bonds are broken

4.1.5 Simple Compressible Systems

In order to demonstrate the meaning and use of the first law with examples and simple calculations, only **simple systems** will be considered. A simple system is one where there is only *one* possible work interaction. The classic example of a simple compressible system is that of a substance contained within a piston-cylinder apparatus, Fig. 4.1.5. The state of the material can be changed either by heat transfer or by the application of work, and the only work interaction possible is the application of a force to the piston head, compressing or expanding the material. Any effects due to magnetic or electrical interactions, or due to motion or gravity, are ignored.

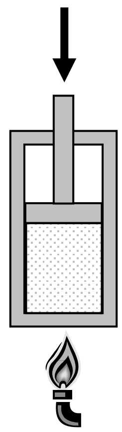


Figure 4.1.5: A simple piston-cylinder system

A **pure substance** is one which has a uniform and invariable chemical composition. In theory this could include different phases of the same substance (e.g. water and steam for H_2O).

In what follows, only pure substances in the context of simple compressible systems will be considered.

Work

If p is the pressure at the piston face, and dV is a small change in volume of the material, Fig. 4.1.5, then the work done in compressing/expanding the material is

$$\delta W = -p\,dV\,, \tag{4.1.6}$$

the minus sign because a positive work is done when the volume gets smaller. The total work done during a compression/expansion of the material is then

$$W = \int \delta W = -\int_{p_1,V_1}^{p_2,V_2} p\,dV \tag{4.1.7}$$

The symbol δ is used here to indicate that the small amount of work δW is not a true differential[4]; it cannot be integrated to a function which is evaluated at only the initial and final states, $W \neq \int_{p_1,V_1}^{p_2,V_2} dW = W(p_2,V_2) - W(p_1,V_1)$, since the work done is process/path. To illustrate this path dependence, consider the $p - V$ graph in Fig. 4.1.6, which shows three different process paths between states 1 (p_1, V_1) and 2 (p_2, V_2). For path ABC, the work done is $p_2(V_2 - V_1)$. For path $\mathrm{AB'C}$, the work done is $p_1(V_2 - V_1)$. The work for the third, curved, path requires an integration along AC and will in general be different from both the other results.

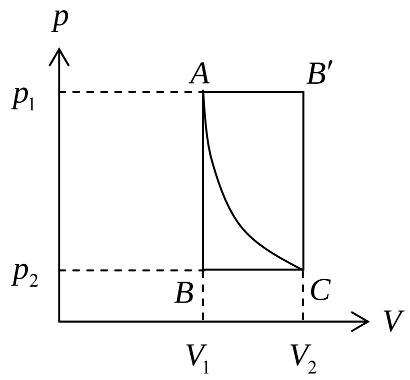


Figure 4.1.6: a p-V diagram

The first law states that $dU = \delta W + \delta Q$ which can now be re-written as

$$\boxed{dU = -p dV + \delta Q} \quad \textbf{First Law for a Simple Compressible System} \quad (4.1.8)$$

4.1.6 Quasi-Static Processes

A system is said to be in **equilibrium** when it experiences no change over time – it is in a **steady state**. Full **thermodynamic equilibrium** of a system requires thermal equilibrium with any surroundings and also mechanical equilibrium[5].

Much of the theory developed here requires that the system be in a certain state with certain properties. If a property such as temperature is varying throughout the material, one cannot easily speak of its "state". Thus when a material is undergoing some process, for example it is being deformed or heated, it is often necessary to assume that it is a **quasi-static** (or **quasi-equilibrium**) process. This means that the process takes place so slowly that the rate of change of the process is slow relative to the time taken for the properties to reach equilibrium. For example, if one heats water in the piston-cylinder arrangement of Fig. 4.1.5 by putting it directly over a hot flame, the water near the base will heat up first and cause convection currents and the water will not be anywhere near an equilibrium state. On the other hand, one could imagine heating the water extremely

[4] but not to be confused with the use of this symbol to represent a variation, as in the context of the principle of virtual work
[5] and also **chemical equilibrium**, where there are no net reactions taking place

slowly with a low flame, so that at any time instant the water temperature is very nearly constant throughout.

To examine what this might mean in the case of the work performed, consider Fig. 4.1.7, which shows the system pressure *p* and the external pressure p_{ext} – the pressure exerted by the surroundings. Assuming thermal equilibrium, if $p = p_{ext}$ then there is full equilibrium. If, however, there is an appreciable difference between the two, for example if a large external pressure is suddenly applied, the piston head will depress rapidly and pressure will not remain uniform throughout the system. However, if the pressures differ by a small amount dp, the work done is

$$W = -\int pdV = -\int (p_{ext} \pm dp)dV = -\int p_{ext}dV \mp \int dpdV = -\int p_{ext}dV \qquad (4.1.9)$$

provided dp is extremely small. The smaller dp, the closer the system will be to mechanical equilibrium. As with the heat transfer, this implies that quasi-equilibrium is maintained provided the piston is moved extremely slowly by incrementally increasing the pressure by very small amounts. (It is often suggested that this might be achieved by repeatedly placing individual grains of sand on the piston head.)

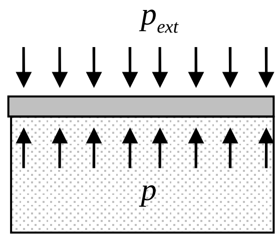


Figure 4.1.7: pressures exerted on a piston head

Unless otherwise stated, it will be assumed that the material at any instance is in quasi-equilibrium. If the system is *not* in equilibrium, Eqn. 4.1.8, $dU = -pdV + \delta Q$, does not make much sense, and one would have to use the more general version $dU = \delta W + \delta Q$.

Example

A gas is contained in a rigid thermally insulated container. It is then allowed to expand into a similar container initially evacuated, Fig. 4.1.8. There is no heat transfer and so $Q = 0$. Since a vacuum provides no resistance to an expanding gas, there is no pressure and hence no work done. Therefore there is no change in the internal energy of the gas. This is *not* a quasi-static process.

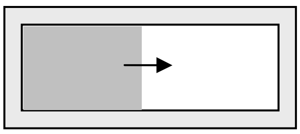


Figure 4.1.8: a thermally insulated gas expanding in an evacuated container

■

Example

Consider the cylinder arrangement of Fig. 4.1.9, which shows a gas contained by a weight. The gas is heated and this causes the weight to rise. The pressure is constant and so the work done is $p\Delta V = p(V_2 - V_1)$. This example shows a system taking heat as input and performing work as output, with no necessary internal energy change.

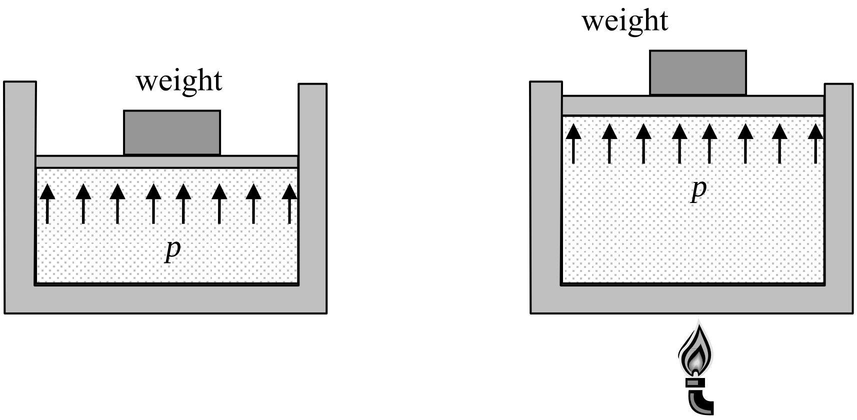


Figure 4.1.9: a heated gas causing a weight to move

■

The opposite process, whereby work is converted purely into heat is called **dissipation** (for example, as can occur in a frictional brake).

4.1.7 State Variables and State Functions

Now a general guide known as the **state principle** says that there is one independent property for each way a system's energy can be varied independently. For a simple compressible system, there are two ways of varying the energy and so the material has *two* **independent properties**[6]. One can take any two of, for example, the temperature[7] θ, pressure p, volume V or internal energy U. The two chosen independent properties are the **state variables** of the system. The state of the system is completely described by these state variables.

Other properties of the system can be expressed as **state functions** of the state variables. For example, suppose that one takes the temperature and volume to be the state variables. Then the relations

$$p = p(\theta, V), \quad U = U(\theta, V) \tag{4.1.10}$$

[6] this is not always the case; it should be emphasised that the number of state variables needed to completely characterise a material undergoing a certain process is, in the final analysis, determined from experiment

[7] the symbol θ denotes the **absolute temperature**, with $\theta > 0$

are state functions for the pressure and internal energy. Equations involving the various properties of a system, as in 4.1.10, are also called **equations of state**. The first of these, relating force variables (in this simple case, the pressure *p*) to kinematic variables (in this case, the volume *V*) and temperature, is called a **thermodynamic** (or **thermal**) **equation of state**. The second, relating the internal energy to a thermal variable (here temperature) and a kinematic variable, is called a **caloric equation of state**.

Different sets of state variables may be chosen. For example, taking p and θ to be the state variables, the state functions would be

$$V = V(p,\theta), \quad U = U(p,\theta) \tag{4.1.11}$$

A key feature of a state function is that its value is determined from the values of the state variables; its value does not depend on the particular path taken to reach the current state. For example, the internal energy is a state function (by its own definition); if one chooses the state variables to be (p,V), the change in internal energy between states '1' and '2' is (compare with Eqn. 4.1.7)

$$\Delta U = \int_{p_1,V_1}^{p_2,V_2} dU = U(p_2,V_2) - U(p_1,V_1) \tag{4.1.12}$$

The value of ΔU depends only on the values of the state variables, in other words its value is the same no matter what path is taken between A and C in Fig. 4.1.6.

$U = U(p,V)$ defines a surface U in $p-V$ space. The total differential of U is then[8]

$$dU = \left(\frac{\partial U}{\partial p}\right)_V dp + \left(\frac{\partial U}{\partial V}\right)_p dV \tag{4.1.13}$$

Although the partial differentiation here means differentiation with respect to one variable only, it is conventional in classical thermodynamics to include a subscript to explicitly indicate this, as here – the subscript emphasises the variable which is held constant. This notation helps avoid confusion when the set of state variables being used is changed during an analysis.

These partial derivatives are themselves state functions; since the function U is known for all (p,V), so are its slopes.

4.1.8 Specific Properties

Specific properties are properties *per unit mass*. They are usually denoted by lower case letters. For example, the specific volume (reciprocal of the density) and specific internal energy are

8

$$v = \frac{V}{m}, \quad u = \frac{U}{m} \tag{4.1.14}$$

where m is the mass of the system. The properties V and U are **extensive properties**, meaning they depend on the amount of substance in the system. The specific properties on the other hand are **intensive properties**, meaning they do not depend on the amount of substance. Other intensive properties are the temperature θ and pressure p.

One can also express the heat and work as per unit mass:

$$\delta q = \frac{\delta Q}{m}, \quad \delta w = \frac{\delta W}{m} \tag{4.1.15}$$

4.1.9 Heat Capacity

Specific Heat and the Enthalpy

The **heat capacity** is defined as *the amount of heat required to raise the system by one unit of temperature*, so the higher the heat capacity, the more the heat required to increase the temperature. For example, water has a very high heat capacity, so it requires a lot of heating to increase its temperature. By the same token, it can give out a lot of heat without dropping in temperature too quickly (hence its use in hot water bottles).

The amount of heat required to raise the temperature by a fixed amount is path-dependent, depending as it does on the amount of accompanying work carried out, so the heat capacity as defined above is ambiguous. To remove this ambiguity, one can specify the path taken during which the heat is added; the two common paths chosen are those at constant volume and at constant pressure.

From Eqn. 4.1.8, the **heat capacity at constant volume** is, by definition,

$$C_V \equiv \left(\frac{\delta Q}{d\theta}\right)_V = \left(\frac{\partial U}{\partial \theta}\right)_V \tag{4.1.16}$$

In this case, all the supplied thermal energy goes into raising the temperature of the body. Note that C_V is a state function; this is clear from the fact that there is no path dependence involved in its evaluation.

The question arises: what is the volume which is held "constant"? Although C_V will in general depend on the V chosen, this dependence is very weak for many materials; a material is usually assigned a value for C_V without reference to the volume at which it is measured.

The **heat capacity at constant pressure** is by definition

$$C_p \equiv \left(\frac{\delta Q}{d\theta}\right)_p = \left(\frac{\partial U}{\partial \theta} + p\frac{\partial V}{\partial \theta}\right)_p = \left(\frac{\partial H}{\partial \theta}\right)_p \tag{4.1.17}$$

where H is the **enthalpy**, defined by

$$H = U + pV \tag{4.1.18}$$

In this case, some of the thermal energy is converted into work, and so $C_p \geq C_V$.

The enthalpy is a state function, since U, p and V are (either state functions or state variables). As with $C_V = (\partial U / \partial \theta)_V$, the heat capacity $C_p = (\partial H / \partial \theta)_p$ is also state function.

Note that, for an incompressible material, $C_V = C_p = \partial U / \partial \theta$ and there is no ambiguity as to its meaning. Most fluids are incompressible, or nearly so, and solids are also often approximated as incompressible for heat capacity measurements. The case of gases will be discussed below.

Internal Energy Measurements

Suppose now that the heat capacity at constant volume has been carefully measured over a given temperature range, by recording the heat required to effect increments in temperature. The internal energy changes within that range can then be found from

$$\Delta U = U_2 - U_1 = \int_{\theta_1}^{\theta_2} C_V d\theta \qquad \text{(constant volume)} \tag{4.1.18}$$

Although this measurement technique requires constant volume processes, since internal energy is a property the results apply to *all* processes.

Some values for the specific internal energy and enthalpy of steam for a range of temperatures, pressures and specific volumes are given in Table 4.1.1 below. The reference state for internal energy (where u is chosen to be zero) is for saturated water at 0.01°C. The corresponding reference state for the enthalpy is obtained from 4.1.17[9].

θ (^0C)	v (m³/kg)	u (kJ/kg)	h (kJ/kg)
120	1.793	2537.3	2716.6
200	2.172	2658.1	2875.3
280	2.546	2779.6	3034.2
360	2.917	2904.2	3195.9

Table 4.1.1a: Properties for steam at pressure $p = 0.1\text{MPa}$

[9] note that u and h can take on negative values, depending on the reference state chosen

p (MPa)	v (m^3/kg)	u (kJ/kg)	h (kJ/kg)
0.035	6.228	2660.4	2878.4
0.100	2.172	2658.1	2875.3
0.300	0.716	2650.7	2865.5
0.500	0.425	2642.9	2855.4

Table 4.1.1b: Properties for steam at temperature $\theta = 200^0\,\mathrm{C}$

4.1.10 The Ideal Gas

A **thermally perfect gas** is one for which the thermal equation of state is

$$pV = mR\theta \quad \text{or} \quad pv = R\theta \tag{4.1.19}$$

where R is the universal gas constant. Further, an **ideal gas** is a thermally perfect gas whose internal energy depends on the temperature only, that is, its caloric equation of state is of the form

$$U = U(\theta) \tag{4.1.20}$$

To justify this expression from a physical point of view, consider a gas at the microscopic level. Internal energy and pressure are related through intermolecular forces. If the pressure is very low, the internal energy is no longer affected by these forces, since the molecules are so far apart, but only by their kinetic energy of motion, i.e. the temperature. Moderate changes in volume will not bring the molecules of gas close enough together to alter this sole dependence on temperature.

When the internal energy is a function of θ *and V*, one has

$$dU = \left(\frac{\partial U}{\partial \theta}\right)_V d\theta + \left(\frac{\partial U}{\partial V}\right)_\theta dV = C_V d\theta + \left(\frac{\partial U}{\partial V}\right)_\theta dV \tag{4.1.21}$$

Thus for an ideal gas

$$dU = C_V d\theta\,. \tag{4.1.22}$$

Example

Consider an ideal gas undergoing a volume change under **isothermal**, i.e. constant temperature, conditions. From 4.1.19, the quantity $pV = p_1V_1 = p_2V_2$ is a constant $mR\theta$. This constrains the process to lie on one particular path in a $p - V$ diagram. Also, from 4.1.22, $dU = 0$ and so $\delta Q = -\delta W$. If an ideal gas expands at constant temperature then the heat input exactly equals the work done against an incrementally changing external pressure.

■

Consider now a process involving work and heat transfer. One has $\delta Q = C_V d\theta + pdV$ and the total heat input is

$$Q = \int \delta Q = \int_{\theta_1}^{\theta_2} C_V(\theta) d\theta + \int_{p_1,V_1}^{p_2,V_2} pdV \tag{4.1.23}$$

The second integral here clearly depends on the exact combination of pressure and volume during the process, so the heat input Q is path dependent, as expected. However, consider the following:

$$\begin{aligned} \int \frac{\delta Q}{\theta} &= \int_{\theta_1}^{\theta_2} \frac{C_V(\theta)}{\theta} d\theta + \int_{p_1,v_1}^{p_2,V_2} \frac{pdV}{\theta} \\ &= \int_{\theta_1}^{\theta_2} \frac{C_V(\theta)}{\theta} d\theta + mR \int_{V_1}^{V_2} \frac{dV}{V} \\ &= \int_{\theta_1}^{\theta_2} \frac{C_V(\theta)}{\theta} d\theta + mR \ln(V_2 / V_1) \end{aligned} \tag{4.1.24}$$

The quantity on the right is now path independent. In fact, for the simple case where C_V is independent of θ, a good approximation for many "near-ideal" gases, one has

$$\int \frac{\delta Q}{\theta} = C_V \ln(\theta_2 / \theta_1) + mR \ln(V_2 / V_1) \tag{4.1.25}$$

This expression means that, *for an ideal gas undergoing a quasi-static process*, although the quantity Q depends on the process, $\int \delta Q / \theta$ does not and so is a property. This property is called the **entropy** of the gas.

4.1.11 Problems

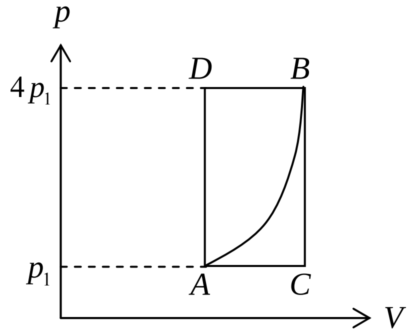
1. A gas is contained in a thermally insulated cylinder. It is very rapidly compressed so that its temperature rises sharply. Has there been a transfer of heat to the gas? Has work been done? Is the process quasi-static?

2. A gas expands from an initial state where $p_1 = 500\,\text{kPa}$ and $V_1 = 0.1\text{m}^3$ to a final state where $p_2 = 100\,\text{kPa}$. The relationship between pressure and volume during the particular process is $pV = k$, a constant. Sketch the process on a $p - V$ diagram and determine the work, in kJ. Interpret the + or – sign on your result.

3. A system, whose equation of state depends only on the volume V, temperature θ and pressure p, is taken (quasi-statically) from state A to state B along the path ACB at the pressures indicated in the figure below. In this process 40J of heat enter the system and 20J of work are done by the system.
   (a) evaluate ΔU
   (b) how much heat enters the system along the path ADB?

(c) if the system goes from B to A by the curved path indicated schematically on the figure, the work done on the system is 30J. How much heat enters or leaves the system?

(d) If the internal energy at A is denoted by U_A, etc., suppose that $U_D - U_A = 20J$. What then is the heat transfer involved in the processes AD and DB?



4. Air is contained in a vertical piston-cylinder assembly by a piston of mass $100\,\text{kg}$ and having a face area of $0.01\,\text{m}^2$. The mass of the air is 5 g, and initially the air occupies a volume of $0.005\ \text{m}^3$. The atmosphere exerts a pressure of $100\,\text{kPa}$ on the top of the piston. Heat transfer of magnitude $2\,\text{kJ}$ occurs slowly from the air to the surroundings, and the volume of the air decreases to $0.002\,\text{m}^3$. Neglecting friction between the piston and the cylinder wall, determine the change in specific internal energy of the air, in kJ/kg. [Note that the pressure is constant on the piston-head, and consists of the piston-weight and the atmospheric pressure.]

5. A closed system, i.e. one which can exchange heat or work with its surroundings, but not matter, undergoes a **thermodynamic cycle**[10] consisting of the following processes:

   Process 1-2: adiabatic compression with $pV^{1.4} = \text{const.}$ from $p_1 = 344.74\,\text{kPa}$, $V_1 = 0.084951\,\text{m}^3$ to $V_2 = V_1/3$

   Process 2-3: constant volume

   Process 3-1: constant pressure, $U_1 - U_3 = 49.27317\,\text{kJ}$

   There are no significant changes in kinetic or gravitational potential energy.

   (a) sketch the cycle on a $p-V$ diagram
   (b) calculate the net work for the cycle
   (c) calculate the heat transfer for process $2-3$

6. How could you use the definition of the specific heat capacity at constant pressure to evaluate the internal energy of a material?

7. Show that for a system (not necessarily an ideal gas) undergoing a constant pressure process, the heat input is equal to the change in enthalpy.

8. Show that, for an ideal gas, $R = C_p - C_V$

[10] meaning the substance is brought back to its initial state at the end of the process; state variables resume their initial values

9. Use the result of problem 8 to show that, when an ideal gas undergoes an adiabatic quasi-static change, $pV^{\gamma} = \text{const.}$ where $\gamma = C_p / C_V$.

10. In Table 4.1.1:
    (a) Does the steam behave like an ideal gas? Nearly? (Note the internal energies in Table 4.1.1b)
    (b) The internal energy decreases as the steam is compressed. Is this what you would expect? Comment.

4.2 Classical Thermodynamics: The Second Law

4.2.1 A Qualitative Sketch of the Second Law and Entropy

The first law of thermodynamics is concerned with the conservation of energy. The second law of thermodynamics is concerned with how that energy is transferred between systems. Its relevance to everyday experience can be seen from the following examples:

- Ice is placed in a glass of water. It melts.
- A hot metal tray is taken out of the oven and placed on a bench top. It cools.
- A brittle plate is dropped from a height onto a hard floor. It smashes into small pieces.
- A piece of iron is left outside. It rusts.
- A bicycle tyre is pumped to high pressure and punctured. The air rushes out.

The common factor in all these examples is that energy is spreading out *in a certain direction*.

- The energy in more rapidly moving warm air molecules disperses to the ice and breaks the intermolecular hydrogen bonds, allowing the water molecules in the ice to move more freely.
- The hot metal contains a relatively large amount of energy due to its vibrating atoms and this energy is transferred to the surrounding air molecules and thereby dispersed.
- The potential energy in the plate disperses through a heating of the surrounding air, the ground and the plate as it smashes.
- The iron atoms and oxygen molecules in the air have chemical (potential) energy stored in their bonds. When iron and oxygen react, lower energy iron oxide bonds are formed and the energy difference is dispersed as heat[1].
- The relatively large energy of the pressurized air in the tyre disperses when the tyre is punctured.

Very qualitatively, the second law says that *energy tends to spontaneously disperse unless hindered from doing so*.

If any of these processes were filmed and the tape accidentally played backwards, the mistake would immediately be evident. However, no physical law (apart from the second law) would be broken if the events happened in reverse. For example, the plate falls because there is a gravitational force pulling it down. However, beginning at the end and working back, it is theoretically possible for the many billions of air (and ground) molecules, which are now moving more rapidly due to the breaking plate, to interact in such a perfect way that the dispersed heat flows back towards the broken pieces and so provides enough energy for the pieces to fly together and gain a kinetic energy to lift off the ground, rise up and eventually slow until it reaches its precise original position off the ground. The second law says that this will not happen; energy does not spontaneously, that is without outside interference, gather together and concentrate in a small locality.

[1] most spontaneous reactions of this type require a certain energy to get started, the **activation energy**, and this hinders the second law from wreaking havoc

Entropy is closely associated with the second law. Again, qualitatively, entropy is *a measure of how dispersed energy is*. Each system has a certain entropy and as energy disperses, the entropy increases. When the air rushes out of the tyre, the entropy of the air and its surroundings increases. When the hot tray cools, the entropy of the tray and surrounding air increases. When the iron and oxygen react, entropy increases.

Entropy can also be defined in terms of probabilities. To take a classic example, consider a box with a shutter splitting it in two. A gas occupies one half of the box (as in Fig. 4.1.8). The shutter is then removed. What will happen? The gas will of course expand to completely fill the box. What are the chances of this happening in reverse? The gas molecules are constantly moving, but the probability of them moving about the box in such a way that all gas molecules would somehow occupy only one half of the box, with no gas in the other half of the box, is zero. As the gas expands, it moves to a more probable state, and the entropy increases.

The Second Law and Maximum Work

When heat is supplied to the confined gas of Fig. 4.1.9, work is done when the gas expands and raises the weight. However, if the flame is not placed under the apparatus but simply left to burn, the heat energy, according to the second law, will disperse into the air. It will not ever *spontaneously* gather back again in a small locality where it could again be used to do some work. The only way to get it back into a small locality again is to input even more energy. In this sense the second law tells us that if we want to maximize the amount of work we can do, we need to use heat energy productively, and if any heat energy escapes it is not possible to use it again without expending more energy. In this sense, entropy can be regarded as a measure of a system's energy unavailable for conversion into work.

A more formal and quantitative treatment of the second law will now be given.

4.2.2 Entropy and the Second Law

Entropy

The entropy S of a system is a property of that system. The change in entropy dS is due to two quantities. First, define the **entropy supply** $\delta S^{(r)}$ (an increment) through

$$\delta S^{(r)} = \frac{\delta Q}{\theta} \tag{4.2.1}$$

where Q is the heat supply; one can imagine the entropy "flowing" into the system. Define also the **entropy production** $\delta S^{(i)}$ (also an increment) to be the difference between the increment of entropy and the entropy supply:

$$dS = \delta S^{(r)} + \delta S^{(i)} \tag{4.2.2}$$

Thus the entropy change in a material is due to two components: the entropy supply, "carried" into the material with the heat supply, and the entropy production, which is produced *within* the material. (The reason for the "*r*" and "*i*" superscripts is given further below.)

Note that, whereas the entropy *S* is a state function (a property), the entropy supply and entropy production are not, since they depend on the particular process by which the state has changed, and hence the use of the symbol "δ" for these functions. (Compare 4.2.2 with the first law, $dU = \delta W + \delta Q$.)

The Second Law

The second law of thermodynamics states that the entropy production is a non-negative quantity,

$$\boxed{\delta S^{(i)} \geq 0} \qquad \textbf{The Second Law} \qquad (4.2.3)$$

Regarding 4.2.2, the Second Law states that the increase in entropy of a system must be at least as great as the entropy flowing into that system.

In terms of the entropy, the first law can be written as

$$\begin{aligned} dU &= \delta W + \theta \delta S^{(r)} \\ &= \delta W + \theta dS - \theta \delta S^{(i)} \end{aligned} \qquad (4.2.4)$$

or, including the second law,

$$\delta W = dU - \theta dS + \theta \delta S^{(i)} \qquad \text{with} \qquad \delta S^{(i)} \geq 0 . \qquad (4.2.5)$$

A process is termed **reversible** if the equality holds, $\delta S^{(i)} = 0$, so that there is no entropy production, in which case $dS = \delta S^{(r)}$. Otherwise it is termed an **irreversible** process, in which case $dS = \delta S^{(r)} + \delta S^{(i)}$. The superscript "*r*" on the entropy supply is to indicate that the entropy supply is equivalent to the change of entropy in a *reversible* process. The superscript "*i*" on the entropy production is to indicate that entropy production is associated with *irreversible* processes.

Alternative Statements of the Second Law

There are many different statements of the second law and each can be "derived" from the others (there is no one agreed version). Another useful definition is that *the heat input to the system in transforming from state A to state B is bounded from above*, according to

$$\delta Q \leq \theta dS \qquad (4.2.6)$$

The maximum possible heat input is θdS, in which case the entropy change is due entirely to entropy supply, with no entropy production – a reversible process. It can be seen that the statement $\delta Q \leq \theta dS$ is equivalent to the statement $\delta S^{(i)} \geq 0$.

A re-arrangement of Eqn. 4.2.6 gives the classic **Clausius's inequality**:

$$dS \geq \frac{\delta Q}{\theta} \tag{4.2.7}$$

4.2.3 Reversibility

Pure Heating

As an example of a reversible process, consider a **pure heating** (or cooling) process, where the volume is held constant, Fig. 4.2.1. Taking the two state variables to be θ and V, it follows from 4.2.5 that the work increment can be expressed as

$$\delta W = \left(\frac{\partial U}{\partial V} - \theta \frac{\partial S}{\partial V} \right)_\theta dV + \left(\frac{\partial U}{\partial \theta} - \theta \frac{\partial S}{\partial \theta} \right)_V d\theta + \theta \delta S^{(i)} \tag{4.2.8}$$

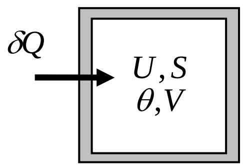


Figure 4.2.1: Pure heating

With $\delta W = dV = 0$, this reduces to

$$\left(\frac{\partial U}{\partial \theta} - \theta \frac{\partial S}{\partial \theta} \right)_V d\theta + \theta \delta S^{(i)} = 0 \tag{4.2.9}$$

Now *U* and *S* are state functions, and the partial derivatives and, in particular, the term inside the brackets are also properties of the system. Further, $\delta S^{(i)} \geq 0$, $\theta > 0$, and $d\theta$ can be positive, negative or zero. Since $d\theta$ can be assigned a value completely independently of the value of the term inside the brackets, the equality in Eqn. 4.2.9 can only be satisfied *in general* if both

$$\left(\frac{\partial U}{\partial \theta} - \theta \frac{\partial S}{\partial \theta} \right)_V = 0 \quad \text{and} \quad \delta S^{(i)} = 0 \tag{4.2.10}$$

The second equality shows that *a quasi-static pure heating process is always reversible*. As mentioned, the first equality is a relation between state functions and is not path-dependent and hence holds for *all* processes, not just for pure heating.

In reality there is never any such thing as a completely reversible process – in the case of pure heating, a reversible process would require that the temperature at any instant is uniform throughout the material, which will never be exactly true. It will be shown below that if there is any appreciable temperature gradient within a material then there will be entropy production.

Reversible Processes

To be precise, a process is reversible when both the system *and its surroundings* can be returned to their original states. For example, if the material in a piston-cylinder arrangement is compressed quasi-statically and there is no friction between the piston and cylinder walls, then the process is reversible – the load can be reduced by very small amounts and the material will "push back" on the piston returning it to its original configuration, with no net work done or heat supplied to the system.

Irreversible Processes

An irreversible process is one for which there is entropy production, $\delta S^{(i)} > 0$. In practice, irreversibilities are introduced into systems whenever there is spontaneity:

- unrestrained expansion of a gas/liquid to a lower pressure – for example when the lid is taken off a gas at high pressure, and is allowed to escape into the atmosphere
- heat transfer from one part of a material to another part at a lower temperature (except in the ideal case where the temperature difference is infinitesimal)
- friction (both the sliding friction of solid on solid and the friction that occurs between molecules in the flow of fluids)

The common factor amongst all these is that the system and its surroundings cannot be returned to their original configurations. For example, with the piston-cylinder arrangement, friction between the piston head and cylinder walls means that further work needs to be expended on the return stroke so that, although the piston-cylinder is returned to its original state, a net amount of work needs to be done and so the "surroundings", or whatever is producing the work, is not back at its original state. Similarly, if the piston was compressed very quickly to its final position, the temperature, momentarily, might well be higher at the piston head than further down in the material. This would produce a spontaneous heat transfer from the upper part of the material to the lower part and it would not be possible to return the system and its surroundings to their original states.

Irreversible Heat Transfer

In the processes studied so far, it has been assumed that all the state functions were uniform throughout the material. In particular, it has been assumed that the temperature is uniform throughout. What if one now has a system whose parts are at different temperatures?

Suppose that a quantity of heat δQ flows from a body at temperature θ_1 to a body at temperature θ_2, Fig. 4.2.2. One can imagine for the sake of argument that the heat capacities of both bodies are sufficiently large that their temperatures are effectively unchanged by the heat flow. The two bodies are insulated from their surroundings.

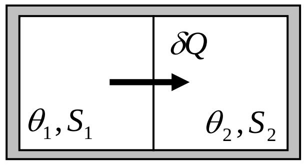


Figure 4.2.2: Heat flow from one body to another

This is pure heating and so the entropy change due to this heat transfer are the entropy supplies $\delta S_1^{(r)} = \delta Q / \theta_1 < 0$ and $\delta S_2^{(r)} = \delta Q / \theta_2 > 0$. Considering now the complete system (both bodies), there is no entropy supply, so any entropy change must be an entropy production

$$\delta S^{(i)} = \frac{\delta Q}{\theta_2} - \frac{\delta Q}{\theta_1} \tag{4.2.11}$$

Since $\delta S^{(i)} \geq 0$, it follows that $\theta_1 > \theta_2$, that is, *heat flows from the warmer body to the colder body*.

In this example there is no work done, no heat transfer and no internal energy change, but there is an entropy change.

If one wants the heat transfer to be very nearly reversible, one can make the entropy production very small. This can be achieved by making the temperature difference between the two bodies very small: by letting $\theta_1 = \theta, \theta_2 = \theta + \Delta\theta$, one has $\delta S^{(i)} \approx -(\delta Q / \theta)(\Delta\theta / \theta)$. Keeping the entropy supply constant, this means that one must make $\Delta\theta / \theta$ as small as possible. Thus heat transfer is reversible *only* if there is an "infinitely small" temperature difference between the two bodies.

Entropy supply is due to heat transfer, but the entropy production here is due to an adiabatic irreversible change.

Entropy Measurements

The entropy of a material can be measured as follows. First, since $dS = \delta Q / \theta + \delta S^{(i)}$, one has $dS = C(d\theta / \theta) + \delta S^{(i)}$ where C is the specific heat capacity. Thus, for a *reversible* process, one has

$$\Delta S = \int_{\theta_1}^{\theta_2} C \, d\theta / \theta \qquad \text{(reversible)} \tag{4.2.12}$$

but one must ensure that the entropy production is zero. In practice, what one does is keep $d\theta / \theta$ small enough so that the entropy production is sufficiently small for the accuracy required. Once the entropy change is found, it of course applies to *all* processes, not just the reversible process used in the experiment.

Thermodynamic Equilibrium

Thermodynamic equilibrium has already been mentioned – it occurs when no changes of the state variables can occur. Thus, one requires that $\delta W = \delta Q = dU = 0$. With $\delta Q = 0$, one has $\delta S^{(r)} = 0$ and $dS = dS^{(i)}$. For full equilibrium, one requires that $dS^{(i)} = 0$ but, since entropy production tends always to increase the entropy, thermal equilibrium can only occur if the entropy has reached its *maximum possible value*.

4.2.4 Free Expansion of an Ideal Gas

It was seen that, for an ideal gas undergoing a quasi-static process (see Eqn. 4.1.25),

$$\Delta S = S_2 - S_1 = C_V \ln(\theta_2 / \theta_1) + mR \ln(V_2 / V_1) \tag{4.2.13}$$

and the entropy production is zero. In other words, any quasi-static process involving an ideal gas is reversible.

Consider an isothermal quasi-static process in which an ideal gas is heated very slowly, so that it does work and expands to twice its original volume. One has $\Delta S = mR \ln 2$. This is the change in entropy in the gas; the surroundings has an equal decrease in entropy as heat leaves it so that, by definition, the total entropy change in this reversible process is zero.

Consider now a thermally insulated container divided by a partition into two parts each of volume *V*. One of these contains an ideal gas and the other is evacuated. The partition is taken away, so that the gas completely fills the container (see Fig. 4.1.8). During the spontaneous expansion, there is a complex non-equilibrium turbulence. The gas eventually settles down to its new equilibrium position. The initial and final states are as for the reversible heating described above, so the entropy change in the gas must again be $\Delta S = mR \ln 2$. However, here there is no interaction with the surroundings, so this entropy must be entropy production.

4.2.5 Problems

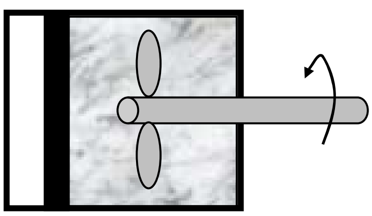
1. A system undergoes a process in which work is done on the system and the heat transfer Q occurs at a constant temperature θ_b. For each case, determine whether the entropy change of the system is positive, negative, zero or indeterminate:
   (a) reversible process, $Q > 0$
   (b) reversible process, $Q = 0$
   (c) reversible process, $Q < 0$
   (d) irreversible process, $Q > 0$
   (e) irreversible process, $Q = 0$
   (f) irreversible process, $Q < 0$

2. A block of lead at temperature $200\ \mathrm{K}$ has heat capacity $C = 1000\ \mathrm{J\,K^{-1}}$, which is independent of temperature in the range $100 - 200\ \mathrm{K}$. It is to be cooled to $100\ \mathrm{K}$ in liquid baths, which are large enough that their temperatures do not change. Assume that the only heat transfer which occurs is between the liquid baths and the lead block, and that the lead block changes temperature instantaneously/uniformly (so that there is no entropy production *within* the block). What is the entropy supply for the lead and the liquid bath(s), and the net entropy production, during the following processes: the lead is
   (a) plunged straight into a liquid bath at $100\ \mathrm{K}$
   (b) first cooled in a bath at $150\ \mathrm{K}$ and then in a second bath at $100\ \mathrm{K}$
   (c) cooled using four baths at temperatures $175\ \mathrm{K}$, $150\ \mathrm{K}$, $125\ \mathrm{K}$ and $100\ \mathrm{K}$
   (d) cooled in an infinite number of temperature baths with a continuous range from $200\ \mathrm{K}$ to $100\ \mathrm{K}$

   [hint: no work is done; use the lead's heat capacity to evaluate Q]

3. Consider an *insulated* piston-cylinder assembly which initially contains water as a saturated liquid at $100\ \mathrm{C}$ (373.15 K), as illustrated below. A paddle wheel acts on the water, which undergoes a process to the corresponding saturated vapour state at the *same* temperature, during which the piston moves freely in the cylinder (no friction). Using the data below, determine
   (a) the net work per unit mass done – which is greater, the work done by the paddle wheel or that done by the expanding water?
   (b) the specific entropy supply; the specific entropy production – why do you think it is non-zero?

   Next, consider the case where the initial and final states are the same as before, but the change is now brought about by the supply of *heat only* (with no paddle wheel). Determine
   (c) the work done per unit mass[2]
   (d) the heat transfer per unit mass
   (e) the specific entropy supply and the specific entropy production – is this a surprise?



	u (kJ / kg)	v (m^3 / kg)	s (kJ / kg.K)	p (MPa)
Liquid	400	0.001	1.448	0.1014
Gas	2500	1.6	7.510	0.1014

[2] the initial and final temperatures and pressures are $100\ \mathrm{C}$ and $0.1014\ \mathrm{MPa}$ – these are the "end-points" for the initial and final *states* – in general, they may not necessarily be constant throughout the process – we do not know (and don't have to know here) how the temperature and pressure changed *during* the process in parts (a-b); with the paddle wheel, the temperature and pressure are unlikely to be uniform throughout the material. For parts (c-e), it is reasonable to assume that they are constant throughout

4. A certain mass of an ideal gas for which $C_V = 3R/2$, independent of temperature, is taken *reversibly* from $\theta = 100$ K, $p = 10^5$ Pa to $\theta = 400$ K, $p = 8\times10^5$ Pa by two different paths (1) and (2):
   (1) consisting of (a) at constant volume from $\theta = 100 \to 400$ K, (b) isothermally to the final pressure
   (2) consisting of (a) at constant pressure from $\theta = 100 \to 400$ K, (b) isothermally to the final volume

   Calculate the entropy changes and verify that the total entropy change is the same for both paths. Compare this with the heat absorbed or given out for each of paths (1) and (2) – they even turn out to be of opposite sign.
   [hint: use the ideal gas law and the fact that for a constant volume process, $\delta Q = C_V d\theta$; also, use Eqn. 4.2.13, the fact that $Q_{\text{rev}} = \int_{S_1}^{S_2} \theta\, dS$, and the result of Q.8 from section 4.1]

4.3 Thermodynamic Functions

Four important and useful thermodynamic functions will be considered in this section (two of them have been encountered in the previous sections). These are the **internal energy** U, the **enthalpy** H, the **Helmholtz free energy** (or simply the **free energy**) Ψ and the **Gibbs free energy** (or simply the **Gibbs function**) G. These functions will be defined and examined below for both reversible and irreversible processes.

4.3.1 Reversible Processes

Consider first a reversible process.

The Internal Energy

The internal energy is

$$\begin{aligned} dU &= \delta W + \delta Q \\ &= -p dV + \theta dS \end{aligned} \tag{4.3.1}$$

the second line being valid for quasi-static processes. The properties of a pure compressible substance include V, θ, S and p. From 4.3.1, it is natural to take V and S as the state variables:

$$dU = \left(\frac{\partial U}{\partial V}\right)_S dV + \left(\frac{\partial U}{\partial S}\right)_V dS \tag{4.3.2}$$

so that

$$p = -\left(\frac{\partial U}{\partial V}\right)_S, \qquad \theta = \left(\frac{\partial U}{\partial S}\right)_V \tag{4.3.3}$$

Thus $U(V,S)$ contains all the thermodynamic information about the system; given V and S one has an expression for U and can evaluate p and θ through differentiation. U is a **thermodynamic potential**, meaning that it provides information through a differentiation.

V and S are said to be the **canonical** (**natural**) **state variables** for U. By contrast, expressing the internal energy as a function of the volume and temperature, for example, $U = U(V,\theta)$, is not so useful, since this cannot provide all the necessary information regarding the state of the material. A new state function will be introduced below which has V and θ as canonical state variables.

Similarly, the equation of state $\theta(V, p)$ does not contain all the thermodynamic information. For example, there is no information about U or S, and this equation of state must be supplemented by another, just as the ideal gas law is supplemented by the caloric equation of state $U = U(\theta)$.

Returning to the internal energy function, and taking the differential relations between p, θ and U, Eqns. 4.3.3, and differentiating them again, and using the fact that $\partial^2 U / \partial V \partial S = \partial^2 U / \partial S \partial V$, one arrives at the **Maxwell relation**,

$$-\left(\frac{\partial p}{\partial S}\right)_V = \left(\frac{\partial \theta}{\partial V}\right)_S \tag{4.3.4}$$

The Helmholtz Free Energy

Define the **(Helmholtz) free energy function** through

$$\Psi = U - \theta S \tag{4.3.5}$$

One has

$$\begin{aligned} d\Psi &= dU - \theta dS - S d\theta \\ &= dU - \delta Q - S d\theta \quad , \\ &= -pdV - Sd\theta \end{aligned} \tag{4.3.6}$$

the second line being valid for reversible processes. Now V and θ have emerged as the natural state variables. Writing $\Psi = \Psi(V, \theta)$,

$$d\Psi = \left(\frac{\partial \Psi}{\partial V}\right)_\theta dV + \left(\frac{\partial \Psi}{\partial \theta}\right)_V d\theta \tag{4.3.7}$$

so that

$$p = -\left(\frac{\partial \Psi}{\partial V}\right)_\theta, \qquad S = -\left(\frac{\partial \Psi}{\partial \theta}\right)_V \tag{4.3.8}$$

The Enthalpy and Gibbs Free Energy

The enthalpy is defined by Eqn. 4.1.18,

$$H = U + pV \tag{4.3.9}$$

To determine the canonical state variables, evaluate the increment:

$$\begin{aligned} dH &= dU + pdV + Vdp \\ &= \delta W + \delta Q + pdV + Vdp \end{aligned} \tag{4.3.10}$$

and so

$$dH = \theta dS + Vdp \tag{4.3.11}$$

and the natural variables are p and S. Finally, the Gibbs free energy function is defined by

$$G = U - \theta S + pV \tag{4.3.12}$$

and the canonical state variables are p and θ.

The definitions, canonical state variables and Maxwell relations for all four functions are summarised in Table 4.3.1 below.

Thermo-dynamic potential	Symbol and appropriate variables	Definition	Differential relationship	Maxwell relation
Internal energy	$U(S,V)$		$dU = -pdV + \theta dS$	$\theta = \left(\frac{\partial U}{\partial S}\right)_V$, $p = -\left(\frac{\partial U}{\partial V}\right)_S$ $\left(\frac{\partial \theta}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V$
Enthalpy	$H(S,p)$	$H = U + pV$	$dH = Vdp + \theta dS$	$\theta = \left(\frac{\partial H}{\partial S}\right)_p$, $V = \left(\frac{\partial H}{\partial p}\right)_S$ $\left(\frac{\partial \theta}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p$
Helmholtz free energy	$\Psi(\theta,V)$	$\Psi = U - \theta S$	$d\Psi = -pdV - Sd\theta$	$S = -\left(\frac{\partial \Psi}{\partial \theta}\right)_V$, $p = -\left(\frac{\partial \Psi}{\partial V}\right)_\theta$ $\left(\frac{\partial S}{\partial V}\right)_\theta = \left(\frac{\partial p}{\partial \theta}\right)_V$
Gibbs free energy	$G(\theta,p)$	$G = U - \theta S + pV$	$dG = Vdp - Sd\theta$	$S = -\left(\frac{\partial G}{\partial \theta}\right)_p$, $V = \left(\frac{\partial G}{\partial p}\right)_\theta$ $\left(\frac{\partial S}{\partial p}\right)_\theta = -\left(\frac{\partial V}{\partial \theta}\right)_p$

Table 4.3.1: Thermodynamic Potential Functions and Maxwell relations

Mechanical variables: whereas the internal energy and the Helmholtz free energy are functions of a kinematic variable (V), the enthalpy and the Gibbs function are functions of a force variable (p).

Thermal variables: whereas the internal energy and the enthalpy are functions of the entropy, the Helmholtz and Gibbs free energy functions are functions of the temperature.

If one is analyzing a process with, for example, *constant* temperature, it makes sense to use either the Helmholtz or Gibbs free energy functions, so that there is only one variable to consider.

Note that the temperature is an observable property and can be controlled to some extent. Values for the entropy, on the other hand, cannot be assigned arbitrary values in experiments. For this reason a description in terms of the free energy, for example, is often more useful than a description in terms of the internal energy.

4.3.2 Irreversible Processes

Consider now an irreversible process.

The Internal Energy

One has $\delta W = dU - \theta dS + \theta \delta S^{(i)}$ and, with the internal energy again a function of the entropy and volume,

$$\delta W = \left[\left(\frac{\partial U}{\partial S} \right)_V - \theta \right] dS + \left(\frac{\partial U}{\partial V} \right)_S dV + \theta \delta S^{(i)} \tag{4.3.13}$$

Consider the case of pure heating $\delta W = dV = \delta S^{(i)} = 0$, so

$$\theta = \left(\frac{\partial U}{\partial S} \right)_V \tag{4.3.14}$$

as in the reversible case. This relation between properties is of course valid for any process, not necessarily a pure heating one. Thus

$$\delta W = \left(\frac{\partial U}{\partial V} \right)_S dV + \theta \delta S^{(i)} \tag{4.3.15}$$

Express the work in the form

$$\boxed{\begin{aligned} \delta W &= dW^{(q)} + \delta W^{(d)} \\ &= A^{(q)} dV + A^{(d)} dV \end{aligned}} \tag{4.3.16}$$

such that the **quasi-conservative force** $A^{(q)}$ is that associated with the work $W^{(q)}$ which is recoverable, whilst the **dissipative force** A^d produces the work $W^{(d)}$ which is dissipated, i.e. associated with irreversibilities.

From 4.3.15,

$$A^{(q)} = \left(\frac{\partial U}{\partial V} \right)_S \tag{4.3.17}$$

and the **dissipative work** $\delta W^{(d)}$ is

$$\delta W^{(d)} = A^{(d)} dV = \theta \delta S^{(i)} \geq 0 \tag{4.3.18}$$

The name *quasi-conservative force* for the $A^{(q)}$ (here, actually a force per area) is in recognition that the internal energy plays the role of a potential in 4.3.17, but it is also a function of the entropy. It can be seen from Eqn. 4.3.17 that the quasi-conservative force is a state function, and equals $-p$ in a fully reversible process.

In the **isentropic** case, $dS = 0$, and one has

$$\delta W = dU + \theta \delta S^{(i)} \tag{4.3.19}$$

This shows that, in the isentropic case, the internal energy is that part of the work which is recoverable.

The Free Energy

Directly from $\Psi = U - \theta S$, with θ and V the independent variables,

$$\frac{\partial \Psi}{\partial \theta} = \left(\frac{\partial U}{\partial \theta} - \theta \frac{\partial S}{\partial \theta} \right) - S, \qquad \frac{\partial \Psi}{\partial V} = \frac{\partial U}{\partial V} - \theta \frac{\partial S}{\partial V} \tag{4.3.20}$$

From the pure heating analysis given earlier, Eqn. 4.2.9-10, the term $\partial U / \partial \theta - \theta \partial S / \partial \theta$ is zero, so

$$S = -\left(\frac{\partial \Psi}{\partial \theta} \right)_V \tag{4.3.21}$$

as in the reversible case and

$$d\Psi = \left(\frac{\partial \Psi}{\partial V} \right)_\theta dV - S d\theta \tag{4.3.22}$$

The work can now be written again as Eqn. 4.3.16, but now with the quasi-conservative force given by {▲Problem 3}

$$A^{(q)} = \left(\frac{\partial \Psi}{\partial V} \right)_\theta \tag{4.3.23}$$

The dissipative work is again given by 4.3.18. Also, $A^{(q)} = -p$ for a reversible process.

In the isothermal case, $d\theta = 0$,

$$\delta W = d\Psi + \theta \delta S^{(i)} \tag{4.3.24}$$

This shows that, in the isothermal case, the free energy is that part of the work which is recoverable.

The quasi-conservative forces for the internal energy and free energy are listed in Table 4.3.2. Note that expressions for quasi-conservative forces are not available in the case of the Enthalpy and Gibbs free energy since they do not permit in their expression increments in volume dV, which are required for expressions of work increment.

Thermo-dynamic potential	Differential relationship	Relations
$U(S,V)$	$dU = -pdV + \theta dS - \theta\delta S^{(i)}$	$\theta = \left(\dfrac{\partial U}{\partial S}\right)_V$, $A^{(q)} = \left(\dfrac{\partial U}{\partial V}\right)_S$
$\Psi(\theta,V) = U - \theta S$	$d\Psi = -pdV - Sd\theta - \theta\delta S^{(i)}$	$S = -\left(\dfrac{\partial \Psi}{\partial \theta}\right)_v$, $A^{(q)} = \left(\dfrac{\partial \Psi}{\partial V}\right)_\theta$

Table 4.3.2: Quasi-Conservative Forces for Irreversible Processes, $\delta W = A^{(q)}dV + \theta\delta S^{(i)} = -pdV$

4.3.3 The Legendre Transformation

The thermodynamic functions can be transformed into one another using a mathematical technique called the **Legendre Transformation**. The Legendre Transformation is discussed in detail in Part IV, where it plays an important role in Plasticity Theory, and other topics. For the present purposes, note that the Legendre transformation of a function $f(x,y)$ is the function $g(\alpha,\beta)$ where

$$g(\alpha,\beta) = \alpha x + \beta y - f(x,y) \tag{4.3.25}$$

and

$$\alpha = \frac{\partial f}{\partial x},\ \beta = \frac{\partial f}{\partial y}, \qquad x = \frac{\partial g}{\partial \alpha},\ y = \frac{\partial g}{\partial \beta} \tag{4.3.26}$$

When only one of the two variables is switched, the transform reads

$$g(\alpha,y) = \alpha x - f(x,y) \tag{4.3.27}$$

where

$$\alpha = \frac{\partial f}{\partial x}, \qquad x = \frac{\partial g}{\partial \alpha} \tag{4.3.28}$$

For example, if one has the function $U(S,V)$ and wants to switch the independent variable from S to θ, Eqn. 4.3.27 leads one to consider the new function

$$g(\theta,V) = \theta S - U(S,V) \tag{4.3.29}$$

and Eqns. 4.3.28 give

$$\theta = \left(\frac{\partial U}{\partial S}\right)_V \quad \text{and} \quad S = \left(\frac{\partial g}{\partial \theta}\right)_V \tag{4.3.30}$$

It can be seen that $g(\theta,V)$ is the negative of the Helmholtz free energy and the two differential relations in 4.3.30 are contained in Table 4.3.1.

4.3.4 Problems

1. By considering reversible processes, derive the differential relationships and the Maxwell relations given in Table 4.3.1 for (a) the enthalpy, (b) the Gibbs free energy

2. Let the two independent variables be V and θ. Consider the internal energy, $U = U(V,\theta)$. Use the pure heating example considered in §4.2 to show that the quasi-conservative force of Eqn. 4.3.17 can also be expressed as
$$A^{(q)} = \left(\frac{\partial U}{\partial V} - \theta \frac{\partial S}{\partial V} \right)_\theta$$

3. Show that Eqn. 4.3.22 leads to Eqn. 4.3.23.

4. Use the Legendre Transformation rule to transform the Helmholtz free energy $\Psi(\theta,V)$ into a function of the variables θ and σ. Derive also the two differential relations analogous to Eqns. 4.3.30. Show that this new function is the negative of the Gibbs energy (use the relation $\Psi = U - \theta S$), where $\sigma = -p$, and that the two differential relations correspond to two of the relations in Table 4.3.1.

5. Use the Legendre Transformation rule to transform the enthalpy $H(S,p)$ into a function of the variables S and V. Show that this new function is the negative of the internal energy, and that the two differential relations correspond to two of the relations in Table 4.3.1.

4.4 Continuum Thermomechanics

The classical thermodynamics is now extended to the thermomechanics of a continuum. The state variables are allowed to vary throughout a material and processes are allowed to be irreversible and move far from thermal and mechanical equilibrium. Some schools of thought would question whether entropy is a state function at all under these conditions. Here, we simply accept the fact that it is. This is part of the **rational thermodynamics** approach and is generally accepted in the solid mechanics community.

4.4.1 The First Law

The first law of thermodynamics is, in rate form,

$$P_{ext} + Q^{*} = \dot{U} + \dot{K} \tag{4.4.1}$$

where P_{ext} is the power of the external forces, Q^{*} is the rate at which heat is supplied (called the **thermal power**, the **non-mechanical power**, or the **rate of thermal work**), $\dot{U}$ is the rate of change of the internal energy and $\dot{K}$ is the rate of change of kinetic energy. The superscript "*" is used here and in what follows to indicate rates of change of quantities which are not state functions.

Recall from Part III, Eqn 3.8.2, the mechanical energy balance,

$$P_{\text{ext}} + P_{\text{int}} = \dot{K} \tag{4.4.2}$$

Eliminating P_{ext} and $\dot{K}$ from these equations leads to

$$-P_{\text{int}} + Q^{*} = \dot{U} \tag{4.4.3}$$

Heat supply

It is convenient to write the total heat supply to a finite volume of material as an integral over the volume. This is done by defining the **heat flux q** to be the rate at which heat is conducted from interior to exterior per unit area, Fig. 4.4.1. The rate of heat *entering* is thus $-\int_s \mathbf{q}\cdot\mathbf{n}ds$. Let there also be a source of heat supply inside the material, for example a radiator of heat. Let $\int_v r dv$ be the rate of such heat supply, where the scalar *r* is the **heat source**, the rate of heat generated per unit volume. Thus, with the divergence theorem,

$$Q^{*} = -\int_v \operatorname{div}\mathbf{q}\, dv + \int_v r dv \tag{4.4.4}$$

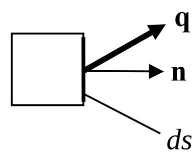


Figure 4.4.1: heat flux vector and normal vector to a surface element

Recall also from Part III, Eqns. 3.8.15, the stress power

$$P_{\text{int}} = -\int_v \boldsymbol{\sigma} : \mathbf{d} dv \tag{4.4.5}$$

Combining Eqns. 4.4.3-5, and expressing the strain energy rate in the form of an integral (see Part III, Eqn. 3.8.15) leads to

$$\int_v \boldsymbol{\sigma} : \mathbf{d} dv - \int_v \text{div}\mathbf{q} dv + \int_v r dv = \int_v \rho \dot{u} dv \tag{4.4.6}$$

Since this holds for all volumes *v*, one has the local form

$$\boxed{\boldsymbol{\sigma} : \mathbf{d} - \text{div}\mathbf{q} + r = \rho \dot{u}} \qquad \textbf{The First Law} \tag{4.4.7}$$

4.4.2 The Second Law

Entropy

The entropy $S(\mathbf{x}, t)$ is defined as the scalar property

$$S = \int_v \rho s dv \tag{4.4.8}$$

where *s* is the specific entropy or **entropy density**. The change in entropy is due to two quantities. First, very like the heat transferred into a body, Eqn. 4.4.4, define the entropy supply $S^{(r)*}$ to be the rate of entropy input,

$$S^{(r)*} = -\int_s \mathbf{s}_\mathbf{q} \cdot \mathbf{n} ds + \int_v s_r dv \tag{4.4.9}$$

where $\mathbf{s}_\mathbf{q}$ is the **entropy flux** through the element surface and s_r is entropy supply due to sources within the element . Further, the entropy flux is taken to be proportional to the heat flux, and the proportionality factor is the reciprocal of the non-negative scalar **absolute temperature** θ (and similarly for the density s_r and the heat supply density *r*) so that, using the divergence theorem,

$$
\begin{aligned}
S^{(r)*} &= -\int_s \frac{\mathbf{q}}{\theta} \cdot \mathbf{n} ds + \int_v \frac{r}{\theta} dv \\
&= -\int_v \operatorname{div}\left(\frac{\mathbf{q}}{\theta}\right) dv + \int_v \frac{r}{\theta} dv
\end{aligned} \tag{4.4.10}
$$

Define the entropy production $S^{(i)*}$ to be the difference between the rate of change of entropy and the entropy supply:

$$
S^{(i)*} = \dot{S} - S^{(r)*} \tag{4.4.11}
$$

The second law of thermodynamics states that the entropy production is a non-negative quantity,

$$
S^{(i)*} \geq 0 \tag{4.4.12}
$$

The Clausius-Duhem Inequality

Thus one has the **Clausius-Duhem inequality**:

$$
S^{(i)*} = \frac{d}{dt} \int_v \rho s dv + \int_v \operatorname{div}\left(\frac{\mathbf{q}}{\theta}\right) dv - \int_v \frac{r}{\theta} dv \geq 0 \tag{4.4.13}
$$

In local form, the Clausius-Duhem inequality reads as (introducing a specific entropy production, $s^{(i)*}$)

$$
s^{(i)*} = \dot{s} + \frac{1}{\rho} \operatorname{div}\left(\frac{\mathbf{q}}{\theta}\right) - \frac{r}{\rho\theta} \geq 0 \tag{4.4.14}
$$

or, equivalently {▲Problem 1},

$$
\boxed{s^{(i)*} = \dot{s} - \frac{r}{\rho\theta} + \frac{1}{\rho\theta} \operatorname{div} \mathbf{q} - \frac{1}{\rho\theta^2} (\mathbf{q} \cdot \nabla\theta) \geq 0} \quad \textbf{The Second Law} \tag{4.4.15}
$$

This is the continuum statement of the Second Law.

4.4.3 The Dissipation Inequality

Eliminating $\operatorname{div} \mathbf{q}$ (and r) from both the first and second laws leads to the **dissipation inequality**

$$
\boxed{\theta s^{(i)*} = \theta\dot{s} - \dot{u} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) - \frac{1}{\rho\theta} (\mathbf{q} \cdot \nabla\theta) \geq 0} \quad \textbf{Dissipation Inequality} \tag{4.4.16}
$$

The term $\theta s^{(i)*}$ is the specific dissipation (or **internal dissipation**) and is denoted by the symbol ϕ. The Clausius-Duhem inequality can simply be written as

$$\phi \equiv \theta s^{(i)*} \geq 0 \tag{4.4.17}$$

Multiplying Eqn. 4.4.16 across by the density leads to

$$\rho\theta s^{(i)*} = \left[\rho\theta\dot{s} - \rho\dot{u} + \boldsymbol{\sigma} : \mathbf{d}\right] + \left[-\frac{1}{\theta}(\mathbf{q} \cdot \nabla\theta)\right] \geq 0 \tag{4.4.18}$$

Each term here has units of power per unit (current) volume. The term inside the first bracket is called the **mechanical dissipation** (per unit volume). The term inside the second bracket is the dissipation due to temperature gradients, i.e. heat flow, and is called the **thermal dissipation** (per unit volume). Note that the thermal dissipation will always be positive if $\mathbf{q}$ and $\nabla\theta$ are of opposite sign. Integrating over a volume v leads to

$$\underbrace{\int_v \rho\theta s^{(i)*} dv}_{\text{dissipation}} = \underbrace{\int_v \left\{\boldsymbol{\sigma} : \mathbf{d} - \rho\dot{u} + \rho\theta\dot{s}\right\} dv}_{\text{mechanical dissipation}} + \underbrace{\int_v \left\{-\frac{1}{\theta}(\nabla\theta \cdot \mathbf{q})\right\} dv}_{\text{thermal dissipation}} \tag{4.4.19}$$

The term $\rho\theta s^{(i)*}$ in Eqn. 4.4.18 is often denoted by the symbol γ and also termed the dissipation. This is a dissipation per unit volume. When the deformations are small, the volume changes are negligible. When the deformations are appreciable, however, the volume and density change, and it is better to work with specific quantities such as ϕ.

The dissipation inequality 4.4.16 is in terms of the internal energy. In terms of the specific free energy $\psi = u - s\theta$, one has

$$\phi = \theta s^{(i)*} = -s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) - \frac{1}{\rho\theta}(\mathbf{q} \cdot \nabla\theta) \geq 0 \tag{4.4.20}$$

4.4.4 The Clausius-Plank Inequality

In many applications the thermal dissipation is very much smaller than the mechanical dissipation (in fact it is zero in many important applications – see section 4.4.7 below). If this is the case then the thermal dissipation rate can be ignored, and one has the stronger form of the second law, in terms of internal energy and free energy,

$$\boxed{\begin{array}{l} \phi = \theta\dot{s} - \dot{u} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) \geq 0 \\ \\ \phi = -s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) \geq 0 \end{array}} \quad \textbf{Clausius-Plank inequality} \qquad (4.4.21)$$

which is known as the **Clausius-Plank inequality**.

Equivalently, one can argue that the processes of mechanical dissipation and heat flow are independent, so that each are separately required to be non-negative, again leading to Eqn. 4.4.21. This issue will be explored more fully in Part IV, where it will indeed be shown that the thermal dissipation is very often decoupled from the mechanical dissipation, with both being required to be separately positive.

Using the first law, Eqn. 4.4.21 can be rewritten in the alternative form

$$\theta\dot{s} = \phi + \frac{1}{\rho}r - \frac{1}{\rho}\mathrm{div}\mathbf{q} \qquad (4.4.22)$$

which is an **evolution equation** for s (showing how it evolves over time).

4.4.5 Special Thermodynamic Processes

Some important limiting processes are considered next.

Reversible Processes

In a reversible process, $\phi = s^{(i)*} = 0$. The Clausius-Plank inequality now becomes

$$\dot{u} = \theta\dot{s} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) \quad \text{or} \quad \dot{\psi} = -s\dot{\theta} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) \qquad (4.4.23)$$

From the First Law,

$$\rho\dot{s} = -\frac{\mathrm{div}\mathbf{q}}{\theta} \qquad (4.4.24)$$

This is the entropy supply (with zero temperature gradients) and corresponds to the classical thermodynamic (for which, also, $\nabla\theta = 0$) expression $dS^{(r)} = \delta Q / \theta$.

Isentropic Conditions

For an isentropic process, the entropy is constant and remains constant, so $\dot{s} = 0$. In this case, the dissipation is

$$\phi = -\dot{u} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) - \frac{1}{\rho\theta}(\mathbf{q} \cdot \nabla\theta) \geq 0 \qquad (4.4.25)$$

Isothermal Conditions

In an isothermal process, the absolute temperature remains constant, $\dot{\theta} = 0$. This can be achieved, for example, by keeping the material's surroundings at constant temperature, and loading the material very slowly, so that any temperature differences which arise between the material and surroundings are allowed to disappear.

The Clausius-Plank inequality becomes

$$\phi = \theta\dot{s} - \dot{u} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) \geq 0 \quad \text{or} \quad \phi = -\dot{\psi} + \frac{1}{\rho}(\boldsymbol{\sigma} : \mathbf{d}) \geq 0 \qquad (4.4.26)$$

Equilibrium Conditions

As mentioned in §4.2.3, a material which is unaffected by external conditions has no work done to it or heat supplied and the first law then states that the internal energy is constant. In that case, when the entropy has reached a maximum and the dissipation is zero, there is no more change in any of the state variables, and equilibrium has been reached.

Adiabatic Conditions

In an adiabatic process, $\mathbf{q} = \mathbf{o}$. This can be achieved, for example, by very rapid loading, so that there is no time for heat exchange with the surroundings.

Under these conditions (and taking also $r = 0$), the first law reads $\boldsymbol{\sigma} : \mathbf{d} = \rho\dot{u}$ (recall that the internal energy change is equal to the work done in an adiabatic process). The dissipation inequality reduces to

$$s^{(i)*} = \dot{s} \geq 0 \qquad (4.4.27)$$

or

$$\phi = \theta\dot{s} \geq 0 \qquad (4.4.28)$$

If the process is both adiabatic and isentropic, then $\phi = \dot{s} = 0$. An adiabatic reversible process is equivalent to an isentropic reversible process.

4.4.6 Summary

A summary of the thermomechanical theory, showing the various laws and relations which are involved, and how they are interconnected, is given in Fig. 4.4.2 below.

Balance of Momentum

$$\mathbf{F} = m\mathbf{a}$$

$$\int_s \mathbf{t}\,\mathrm{d}s + \int_v \mathbf{b}\,\mathrm{d}v = \int_v \rho\dot{\mathbf{v}}\,\mathrm{d}v$$

$$\mathrm{div}\boldsymbol{\sigma} + \mathbf{b} = \rho\dot{\mathbf{v}}$$

$$\Leftrightarrow$$

Work-Energy Principle

$$W_{\mathrm{ext}} + W_{\mathrm{int}} = K$$

$$P_{\mathrm{ext}} + P_{\mathrm{int}} = \dot{K}$$

$$\int_s \mathbf{t}\cdot\mathbf{v}\,\mathrm{d}s + \int_v \mathbf{b}\cdot\mathbf{v}\,\mathrm{d}v - \int_v \boldsymbol{\sigma}:\mathbf{d}\,\mathrm{d}v = \int_v \rho\mathbf{v}\cdot\dot{\mathbf{v}}\,\mathrm{d}v$$

First Law

$$W + Q = U + K$$

$$P_{\mathrm{ext}} + Q^* = \dot{U} + \dot{K}$$

$$\int_s \mathbf{t}\cdot\mathbf{v}\,\mathrm{d}s + \int_v \mathbf{b}\cdot\mathbf{v}\,\mathrm{d}v - \int_s \mathbf{q}\cdot\mathbf{n}\,\mathrm{d}s + \int_v r\,\mathrm{d}v = \int_v \rho\dot{u}\,\mathrm{d}v + \int_v \rho\mathbf{v}\cdot\dot{\mathbf{v}}\,\mathrm{d}v$$

$$\int_v \boldsymbol{\sigma}:\mathbf{d}\,\mathrm{d}v - \int_v \mathrm{div}\mathbf{q}\,\mathrm{d}v + \int_v r\,\mathrm{d}v = \int_v \rho\dot{u}\,\mathrm{d}v$$

$$\boldsymbol{\sigma}:\mathbf{d} - \mathrm{div}\mathbf{q} + r = \rho\dot{u}$$

Second Law

$$S^{(i)} = S - S^{(r)} \ge 0$$

$$S^{(i)*} = \dot{S} - S^{(r)*} \ge 0$$

$$\int_v \rho s^{(i)*}\,\mathrm{d}v = \int_v \rho\dot{s}\,\mathrm{d}v + \int_s \frac{\mathbf{q}}{\theta}\cdot\mathbf{n}\,\mathrm{d}s - \int_v \frac{r}{\theta}\mathrm{d}v \ge 0$$

$$\rho s^{(i)*} = \rho\dot{s} + \frac{1}{\theta}\mathrm{div}\mathbf{q} - \frac{1}{\theta^2}\mathbf{q}\cdot\nabla\theta - \frac{r}{\theta} \ge 0$$

Clausius-Duhem (Dissipation) Inequality

$$\gamma = \rho\phi = \rho\theta s^{(i)*} = \rho\theta\dot{s} - \rho\dot{u} + \boldsymbol{\sigma}:\mathbf{d} - \frac{1}{\theta}\mathbf{q}\cdot\nabla\theta \ge 0$$

$$\gamma = \rho\phi = \rho\theta s^{(i)*} = -\rho s\dot{\theta} - \rho\dot{\psi} + \boldsymbol{\sigma}:\mathbf{d} - \frac{1}{\theta}\mathbf{q}\cdot\nabla\theta \ge 0$$

Clausius-Planck Inequality

$$\phi = \theta s^{(i)*} = \theta\dot{s} - \dot{u} + \frac{1}{\rho}\boldsymbol{\sigma}:\mathbf{d} \ge 0$$

$$\phi = \theta s^{(i)*} = -s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}\boldsymbol{\sigma}:\mathbf{d} \ge 0$$

Figure 4.4.2: Thermomechanics

4.4.7 The use of Thermomechanics in Developing Material Models

Material models and the consequences of the laws of thermomechanics will be discussed in depth in Part IV. Here, as an introduction, the case of small-strain elastic/thermoelastic materials will be touched upon briefly.

The classical thermodynamic expression for the work done to a system is $\delta W = -pdV$. This was generalised to the continuum statement for the power exerted on an infinitesimal element, $\boldsymbol{\sigma} : \mathbf{d}$. In the same way, the kinematic variable of the classical thermodynamic system, *V*, is generalised to the case of a continuum by considering the state to be a function of the *strains*.

When the strains are small, the rate of deformation is equivalent to the time rate of change of the small strain tensor:

$$\mathbf{d} = \dot{\boldsymbol{\varepsilon}}, \qquad d_{ij} = \dot{\varepsilon}_{ij} \tag{4.4.29}$$

The dissipation inequalities are then

$$\phi = \theta\dot{s} - \dot{u} + \frac{1}{\rho}\sigma_{ij}\dot{\varepsilon}_{ij} - \frac{1}{\rho\theta}q_i\theta_{,i} \ge 0, \qquad \phi = \theta\dot{s} - \dot{u} + \frac{1}{\rho}\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \frac{1}{\rho\theta}\mathbf{q}\cdot\nabla\theta \ge 0$$

or

$$\phi = -s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}\sigma_{ij}\dot{\varepsilon}_{ij} - \frac{1}{\rho\theta}q_i\theta_{,i} \ge 0, \quad \phi = -s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \frac{1}{\rho\theta}\mathbf{q}\cdot\nabla\theta \ge 0 \tag{4.4.30}$$

Reversible Process

In the case of reversible processes:

$$\theta\dot{s} - \dot{u} + \frac{1}{\rho}\sigma_{ij}\dot{\varepsilon}_{ij} - \frac{1}{\rho\theta}q_i\theta_{,i} = 0, \qquad \theta\dot{s} - \dot{u} + \frac{1}{\rho}\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \frac{1}{\rho\theta}\mathbf{q}\cdot\nabla\theta = 0$$

or

$$-s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}\sigma_{ij}\dot{\varepsilon}_{ij} - \frac{1}{\rho\theta}q_i\theta_{,i} = 0, \qquad -s\dot{\theta} - \dot{\psi} + \frac{1}{\rho}\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \frac{1}{\rho\theta}\mathbf{q}\cdot\nabla\theta = 0 \tag{4.4.31}$$

Consider now a material whose state is completely defined by the set of state variables $(\theta, \varepsilon_{ij})$, so the free energy state function is $\psi = \psi(\theta, \varepsilon_{ij})$. This defines the **thermoelastic material**. The internal energy is expressed in canonical form as $u = u(s, \varepsilon_{ij})$ (see section 4.3). One then has

$$\dot{u} = \frac{\partial u}{\partial s}\dot{s} + \frac{\partial u}{\partial \varepsilon_{ij}}\dot{\varepsilon}_{ij} \qquad \text{and} \qquad \dot{\psi} = \frac{\partial \psi}{\partial \theta}\dot{\theta} + \frac{\partial \psi}{\partial \varepsilon_{ij}}\dot{\varepsilon}_{ij} \tag{4.4.32}$$

Thus

$$\left(\theta - \frac{\partial u}{\partial s}\right)\dot{s} + \left(\frac{1}{\rho}\sigma_{ij} - \frac{\partial u}{\partial \varepsilon_{ij}}\right)\dot{\varepsilon}_{ij} + \frac{1}{\rho\theta} q_i \theta_{,i} = 0$$
$$\left(\theta - \frac{\partial u}{\partial s}\right)\dot{s} + \left(\frac{1}{\rho}\boldsymbol{\sigma} - \frac{\partial u}{\partial \boldsymbol{\varepsilon}}\right) : \dot{\boldsymbol{\varepsilon}} + \frac{1}{\rho\theta}\mathbf{q}\cdot\nabla\theta = 0$$
or
$$-\left(s + \frac{\partial \psi}{\partial \theta}\right)\dot{\theta} + \left(\frac{1}{\rho}\sigma_{ij} - \frac{\partial \psi}{\partial \varepsilon_{ij}}\right)\dot{\varepsilon}_{ij} - \frac{1}{\rho\theta} q_i \theta_{,i} = 0$$
$$-\left(s + \frac{\partial \psi}{\partial \theta}\right)\dot{\theta} + \left(\frac{1}{\rho}\boldsymbol{\sigma} - \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}}\right) : \dot{\boldsymbol{\varepsilon}} - \frac{1}{\rho\theta}\mathbf{q}\cdot\nabla\theta = 0$$

(4.4.33)

Consider now the free energy equation here (the argument which follows applies also to the internal energy equation). The state variables are $(\theta, \varepsilon_{ij})$ and the other properties are state functions of these variables; these include all the terms inside the brackets, that is, s, ψ, σ_{ij} and also the partial derivatives. For any particular state, these properties will have certain values. On the other hand, no matter the state, the variables $\dot{\theta}, \dot{\varepsilon}_{ij}, \theta_{,i}$ can (in theory) be assigned arbitrary values: negative, zero or positive. The terms which pre-multiply $\dot{\theta}, \dot{\varepsilon}_{ij}$ are *completely independent* of these variables – indeed, this fact is built into the model under consideration: the state is a function of $(\theta, \varepsilon_{ij})$ only and is not, for example, dependent on the values of $\dot{\theta}, \dot{\varepsilon}_{ij}, \theta_{,i}$. On the other hand, $\mathbf{q}$ is not a state function and in fact may be a function of the temperature gradients. The only way that Eqn. 4.4.33 can be satisfied in general, then, is for

$$\nabla\theta = 0 \qquad (4.4.34)$$

and the following relations must hold (compare these with 4.3.3 and 4.3.8):

$$\boxed{\theta = \frac{\partial u}{\partial s}, \quad \sigma_{ij} = \rho\frac{\partial u}{\partial \varepsilon_{ij}}} \quad \text{and} \quad \boxed{s = -\frac{\partial \psi}{\partial \theta}, \quad \sigma_{ij} = \rho\frac{\partial \psi}{\partial \varepsilon_{ij}}} \qquad (4.4.35)$$

Constitutive Relations for Small-Strain Reversible Processes

Note that the density (and volume) changes for small strains may be neglected, so that the density in 4.4.35 can be taken to be the current density or the density in the undeformed configuration, ρ_0.

4.4.8 Thermomechanics in the Material Form

The dissipation inequality is derived here for the case of the material description.

The First Law

In order to rewrite the energy balance equations in material form, first introduce the scalars (what follows is analogous to the definitions of traction and stress with respect to the current and reference configurations)

$$\begin{aligned} q^{(\mathbf{n})} &= -\mathbf{q}\cdot\mathbf{n} \\ Q^{(\mathbf{N})} &= -\mathbf{Q}\cdot\mathbf{N} \end{aligned} \tag{4.4.36}$$

Here $\mathbf{q}$ is the **Cauchy heat flux** of Eqn. 4.4.4, defined per unit current surface area ds with outward normal $\mathbf{n}$, and $\mathbf{Q}$ the **Piola-Kirchhoff heat flux**, defined per unit reference surface area dS and outward normal $\mathbf{N}$.

The rate of heat transfer into the material can now be written as either of

$$-\int_s \mathbf{q}\cdot\mathbf{n}ds = -\int_S \mathbf{Q}\cdot\mathbf{N}dS \tag{4.4.37}$$

Using Nanson's formula, Part III, Eqn. 2.2.59, $\mathbf{n}ds = J\mathbf{F}^{-\mathrm{T}}\mathbf{N}dS$, the Cauchy and Piola-Kirchhoff heat flux vectors are related through $\mathbf{Q} = J\mathbf{F}^{-1}\mathbf{q}$.

The combination of the mechanical energy balance with the first law, i.e. Eqn. 4.4.3, then reads (see also Part III, Eqn. 3.7.26)

$$\int_V \mathbf{P}:\dot{\mathbf{F}}dV - \int_V \mathrm{Div}\mathbf{Q}dV + \int_V RdV = \int_V \rho_0\dot{\mathrm{u}}dV \tag{4.4.38}$$

where $\int_V RdV = \int_v rdv$, or, in local form,

$$\mathbf{P}:\dot{\mathbf{F}} - \mathrm{Div}\mathbf{Q} + R = \rho_0\dot{\mathrm{u}} \tag{4.4.39}$$

or

$$\mathbf{S}:\dot{\mathbf{E}} - \mathrm{Div}\mathbf{Q} + R = \rho_0\dot{\mathrm{u}} \tag{4.4.40}$$

Note that, comparing the spatial and material forms,

$$\int_V \mathrm{Div}\mathbf{Q}dV = \int_v \mathrm{div}\mathbf{q}dv, \quad \mathrm{Div}\mathbf{Q} = J\mathrm{div}\mathbf{q}, \quad \frac{1}{\rho_0}\mathrm{Div}\mathbf{Q} = \frac{1}{\rho}\mathrm{div}\mathbf{q} \tag{4.4.41}$$

The Second Law

Analogous to Eqn. 4.4.13, the second law can be expressed in material form as

$$\frac{d}{dt}\int_V \rho_0 s dV + \int_V \mathrm{Div}\left(\frac{\mathbf{Q}}{\theta}\right) dV - \int_V \frac{R}{\theta} dV \geq 0 \tag{4.4.42}$$

or, analogous to 4.4.16, one has the dissipation inequality

$$\theta \dot{s}^{(i)*} = \theta \dot{s} - \dot{u} + \frac{1}{\rho_0}\left(\mathbf{P} : \dot{\mathbf{F}}\right) - \frac{1}{\rho_0 \theta}(\mathbf{Q} \cdot \mathrm{Grad}\theta) \geq 0 \tag{4.4.43}$$

Isothermal Conditions

In an isothermal process,

$$\phi = \theta \dot{s} - \dot{u} + \frac{1}{\rho_0}\left(\mathbf{P} : \dot{\mathbf{F}}\right) \geq 0 \quad \text{or} \quad \phi = -\dot{\psi} + \frac{1}{\rho_0}\left(\mathbf{P} : \dot{\mathbf{F}}\right) \geq 0 \tag{4.4.44}$$

The second of these can be written as

$$\mathbf{P} : \dot{\mathbf{F}} = \rho_0 \dot{\psi} + \rho_0 \phi \qquad \text{with} \qquad \phi \geq 0 \tag{4.4.45}$$

where the rate of free energy $\left(\rho_0 \dot{\psi}\right)$ and the dissipation $\left(\rho_0 \phi\right)$ are per unit reference volume.

4.4.9 Objectivity

By definition, the scalars heat Q, internal energy U, entropy S and temperature θ are objective, that is they remain unchanged under an observer transformation 2.8.7. It follows that the heat flux vector $\mathbf{q}$ is also objective, transforming according to 2.8.10. By definition, the vector entropy flux $\mathbf{s}_\mathbf{q}$ is objective, that is it transforms according to 2.8.10.

4.4.10 Problems

1. Show that $\mathrm{div}\left(\frac{\mathbf{q}}{\theta}\right) = \frac{1}{\theta}\mathrm{div}\mathbf{q} - \frac{1}{\theta^2}(\mathbf{q} \cdot \nabla \theta)$

2. Show that the relation $\mathbf{Q} = J\mathbf{F}^{-1}\mathbf{q}$ is consistent with the relation 4.4.41, $\mathrm{Div}\mathbf{Q} = J\mathrm{div}\mathbf{q}$.