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

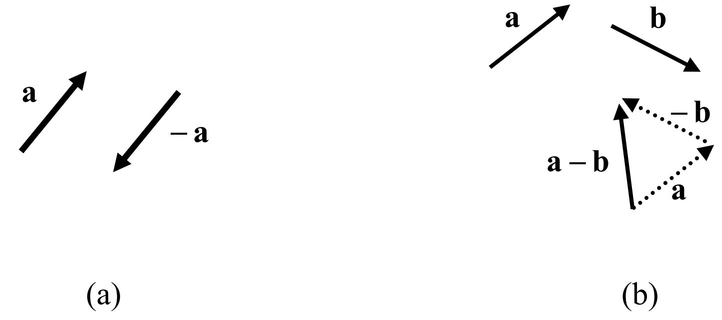


Figure 1.1.2: (a) negative of a vector; (b) subtraction of vectors

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.

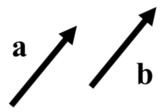


Figure 1.1.3: equal vectors

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.

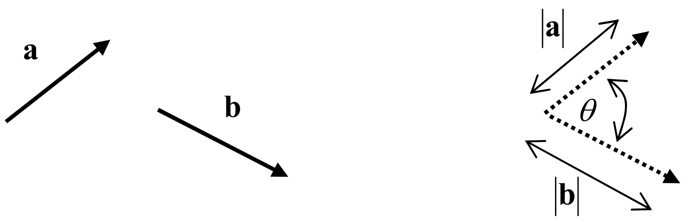


Figure 1.1.4: the dot product

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.

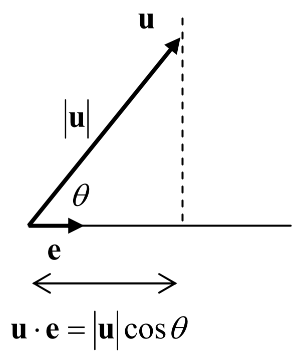


Figure 1.1.5: the projection of a vector along the direction of a unit vector

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

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

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

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta \qquad (1.1.3)$$

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**.

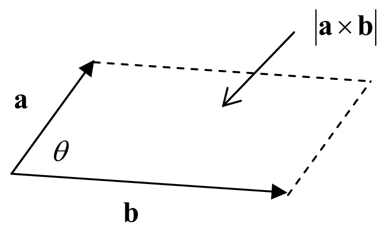


Figure 1.1.6: the magnitude of the cross product

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.

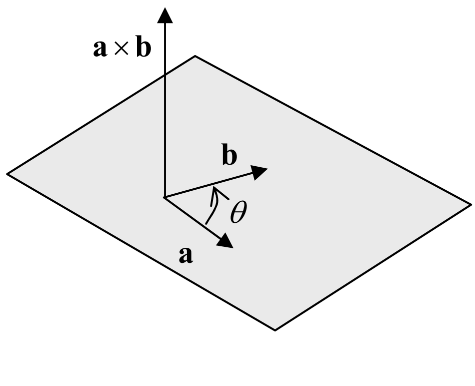


Figure 1.1.7: examples of the cross product

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

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

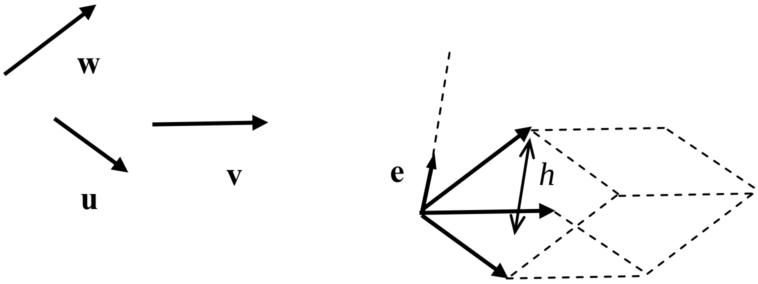


Figure 1.1.8: the triple scalar product

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.

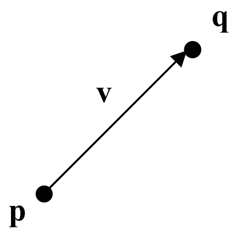
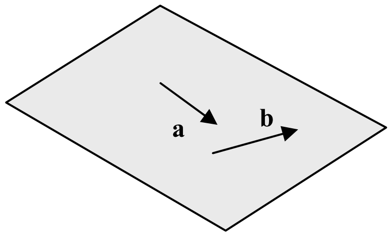


Figure 1.1.9: adding vectors to points

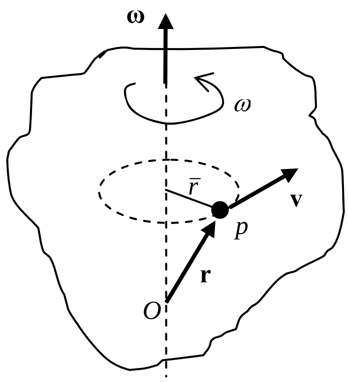
1.1.7 Problems

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

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_2 \cdot \mathbf{e}_3 = \mathbf{e}_3 \cdot \mathbf{e}_1 = 0, \tag{1.3.1}$$

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

$$\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_3 \cdot \mathbf{e}_3 = 1, \tag{1.3.2}$$

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

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

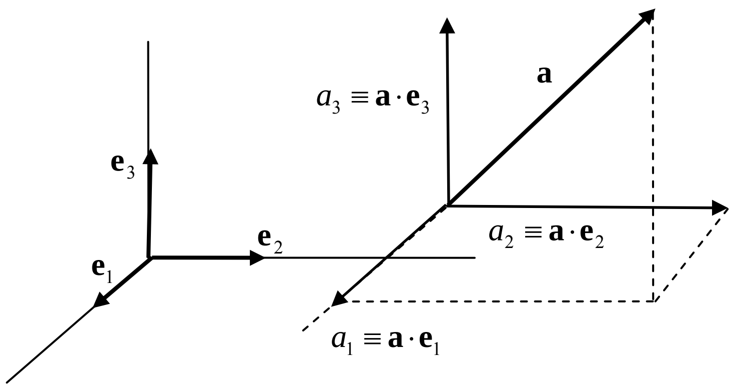


Figure 1.3.1: an orthonormal set of base vectors and Cartesian components

[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

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

Similarly, the cross product is

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

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

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

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

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

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:

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

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

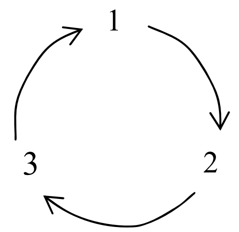


Figure 1.3.2: schematic for the permutation symbol (clockwise gives +1)

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}

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

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

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

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.

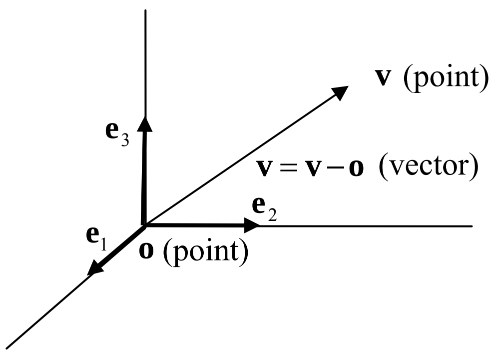


Figure 1.3.3: a Cartesian coordinate system

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

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

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

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

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.

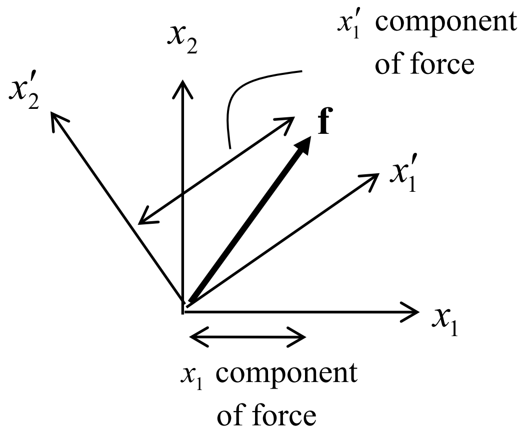


Figure 1.5.1: a vector represented using two different coordinate systems

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".

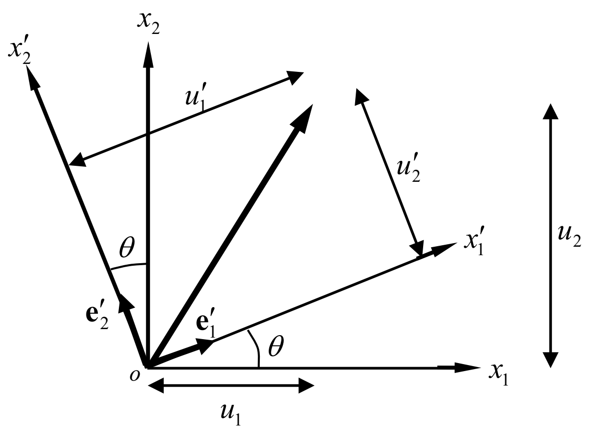


Figure 1.5.2: a vector represented using two different coordinate systems

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

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

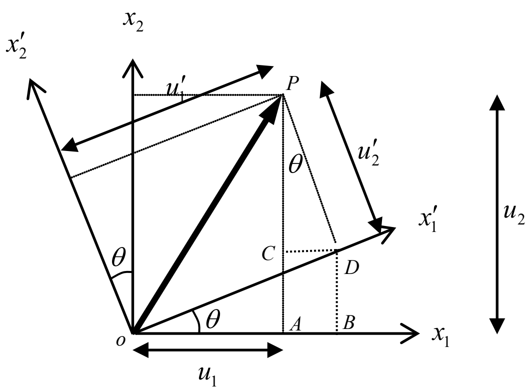


Figure 1.5.3: geometry of the 2D coordinate transformation

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

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

Three Dimensions

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

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

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:

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

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

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

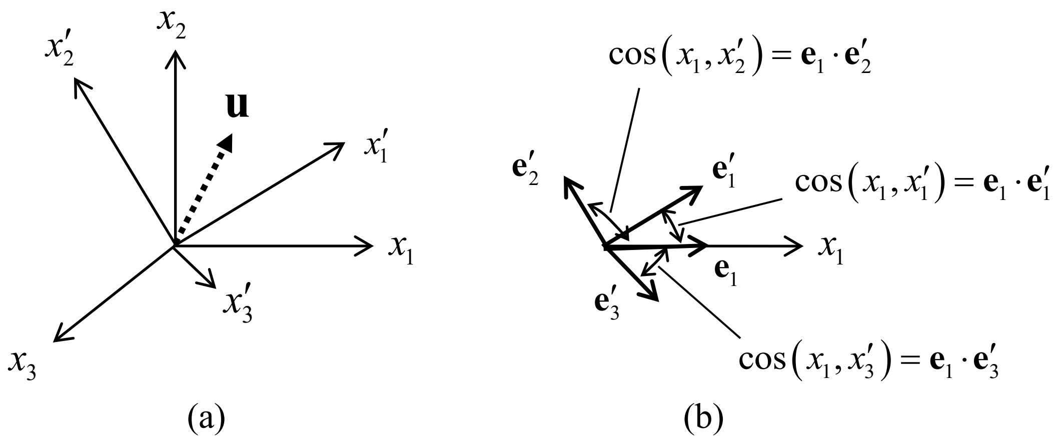


Figure 1.5.4: a 3D space: (a) two different coordinate systems, (b) direction cosines

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

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

One has the general 3D transformation matrix equations

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

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)):

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

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 $\boldsymbol{\tau}$:

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

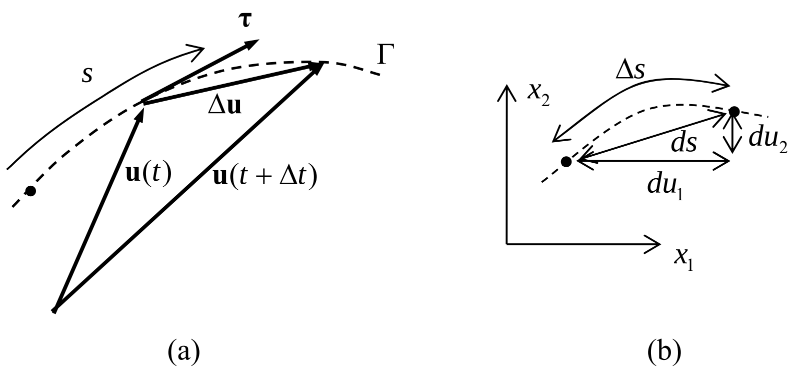


Figure 1.6.1: a space curve; (a) the tangent vector, (b) increment in arc length

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 $d(\boldsymbol{\tau}\cdot\boldsymbol{\tau})/ds = 0$, but also,

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

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

$$\mathbf{v} = \frac{1}{\kappa}\frac{d\boldsymbol{\tau}}{ds} \qquad (1.6.5)$$

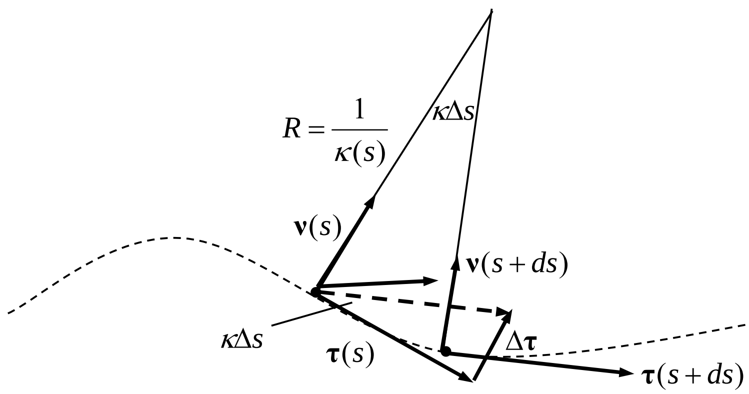


Figure 1.6.2: the curvature

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**:

$$\mathbf{b} = \boldsymbol{\tau} \times \mathbf{v} \qquad (1.6.6)$$

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**.

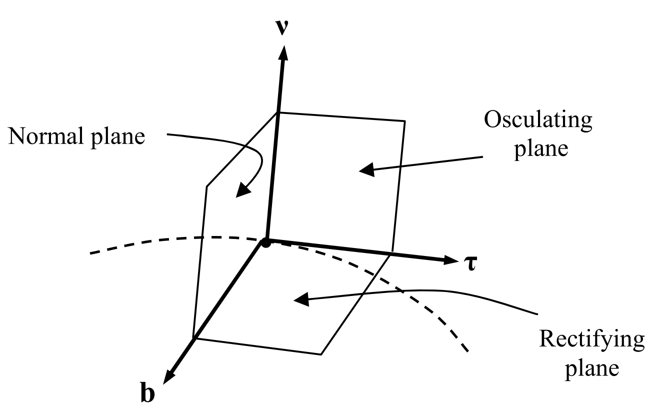


Figure 1.6.3: the unit tangent, principal normal and binormal vectors and associated planes

Rules of Differentiation

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

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

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 $\mathbf{x}$. 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)

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

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:

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

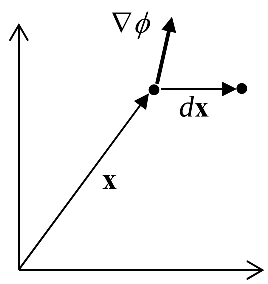


Figure 1.6.4: the gradient of a vector

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 $\theta =$ 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 $\nabla\theta$

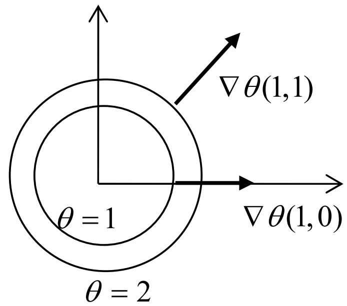


Figure 1.6.5: gradient of a temperature field

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 $\text{grad} f \cdot (d\mathbf{r}/dt) = 0$ and, as seen in §1.6.2, $d\mathbf{r}/dt$ is a vector tangential to the surface. Thus $\text{grad} f$ is normal to the tangent vector; $\text{grad} f$ 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

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

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

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

The following terminology is used:

$$\begin{aligned} \text{grad}\, \phi &= \nabla \phi \\ \text{div}\, \mathbf{u} &= \nabla \cdot \mathbf{u} \\ \text{curl}\, \mathbf{u} &= \nabla \times \mathbf{u} \end{aligned} \qquad (1.6.16)$$

[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

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

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

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

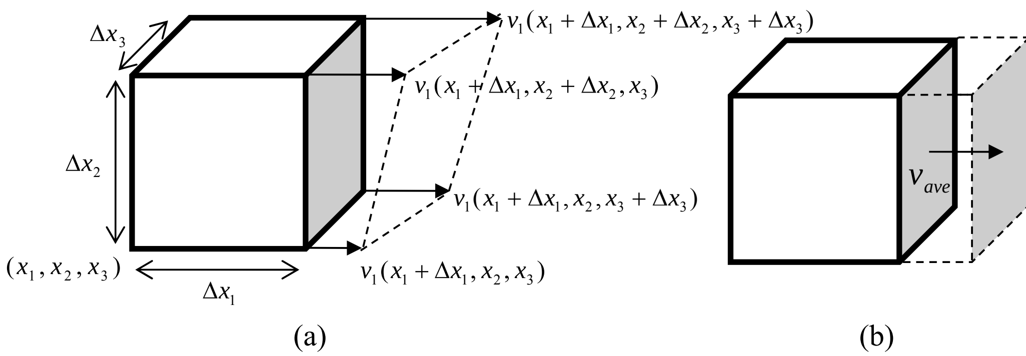


Figure 1.6.6: a differential element; (a) flow through a face, (b) volume of material flowing through the face

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

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

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 **v** for which $\operatorname{div}\mathbf{v} = 0$ is said to be **solenoidal**.

Notes:

- 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 **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

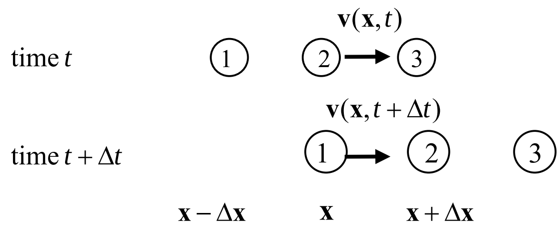


Figure 1.6.7: moving material particles

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

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

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

$$\operatorname{div}\mathbf{q} = r \qquad (1.6.19)$$

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), $\operatorname{div}\mathbf{q} = r$, and assuming the conductivity is constant, leads to $-k\,\nabla\cdot\nabla\theta = r$. Now

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

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

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

It can also be expressed in the form

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

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 $\operatorname{curl}\mathbf{a}$ 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}

$$\operatorname{curl}\mathbf{v} = \nabla \times \left(\boldsymbol{\omega} \times \mathbf{r}\right) = 2\boldsymbol{\omega} \tag{1.6.24}$$

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 $\operatorname{curl}\mathbf{v} \neq 0$, in which case the velocity field **v** is called a **vortex**

field. The paddle wheel would remain stationary in regions where $\text{curl}\,\mathbf{v} = 0$, 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}:

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

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

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

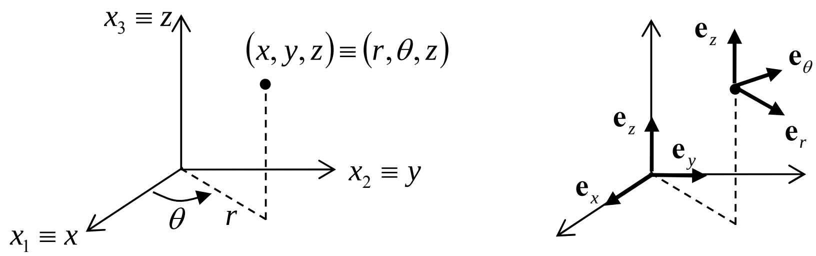


Figure 1.6.8: cylindrical coordinates

The base vectors are related through

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

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:

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

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

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

Spherical Coordinates

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

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

and the base vectors are related through

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

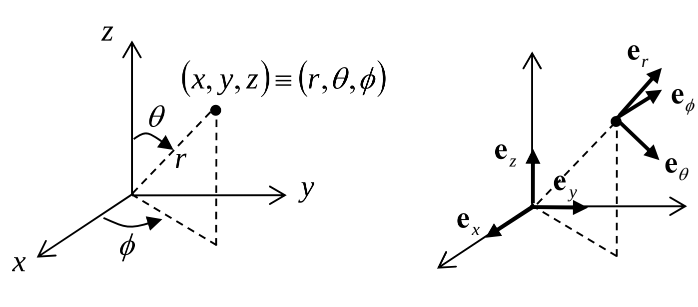


Figure 1.6.9: spherical coordinates

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

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

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

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

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,

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

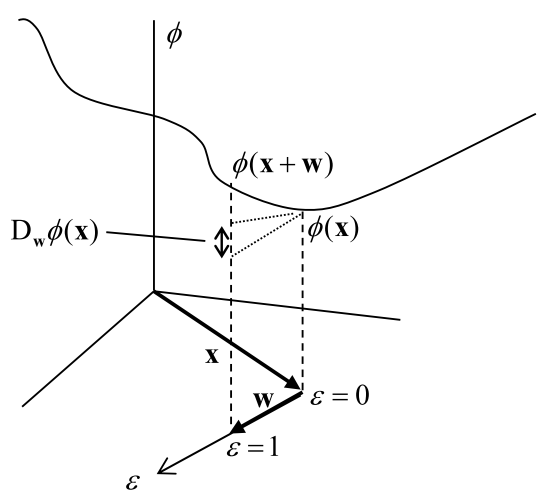


Figure 1.6.10: the directional derivative

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 **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 **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 **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 **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?

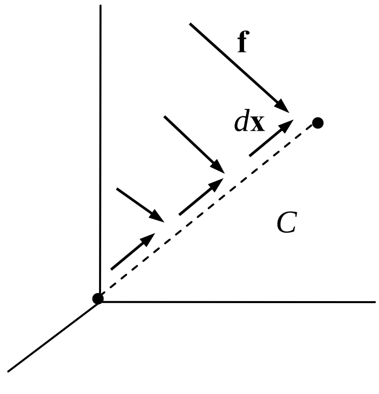


Figure 1.7.1: a particle moving in a force field

Solution
The work done is

$$W = \int_C \mathbf{f} \cdot d\mathbf{x} = \int_C \left\{ \left(3x_1^2 + 6x_2\right) dx_1 - 14x_2 x_3 dx_2 + 20x_1 x_3^2 dx_3 \right\}$$

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, $\text{curl}\mathbf{f} = \mathbf{o}$, and the gravitational scalar potential is $\phi = -mgx_3$:

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

■

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

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

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

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

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

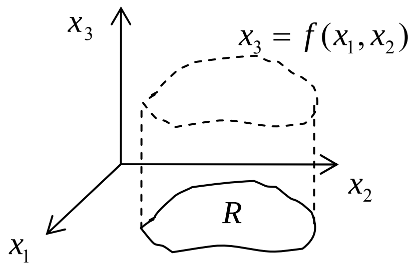


Figure 1.7.2: integration over a region

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)

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

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

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

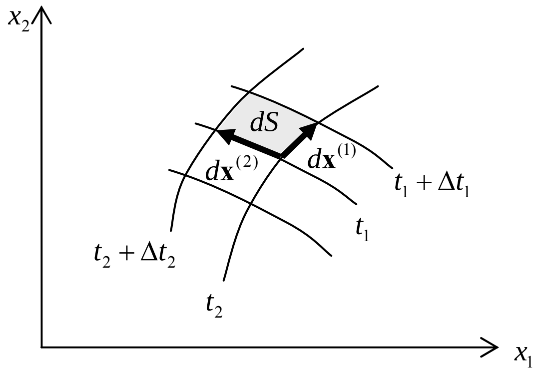


Figure 1.7.3: a surface element

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,

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

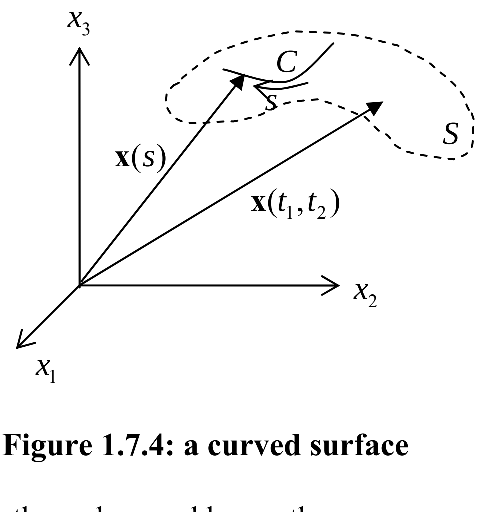


Figure 1.7.4: a curved surface

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 $\mathbf{n} = \operatorname{grad} g / |\operatorname{grad} g|$.

Example (Parametric Representation and the Normal to a Sphere)

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

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

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.

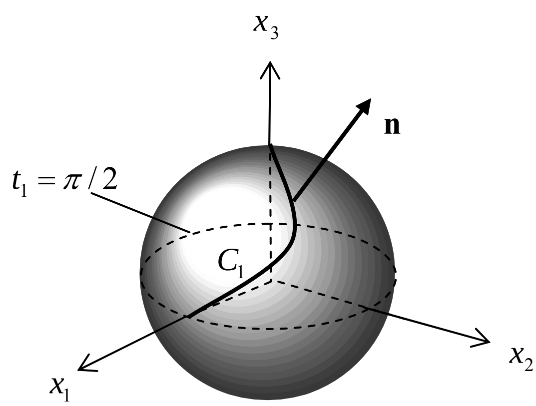


Figure 1.7.5: a sphere

The partial derivatives with respect to the parameters are

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

so that

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

[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.

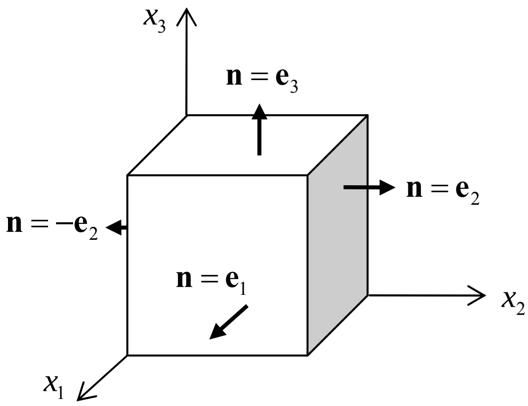


Figure 1.7.6: the unit cube

<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

$$\iint_S \mathbf{f}\cdot\mathbf{n}dS = \int_0^1\int_0^1 \left(4x_3\mathbf{e}_1 - x_2^2\mathbf{e}_2 + x_2x_3\mathbf{e}_3\right)\cdot\mathbf{e}_1 dx_2 dx_3 = 4\int_0^1\int_0^1 x_3 dx_2 dx_3 = 2$$

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: $\mathbf{v}\cdot\mathbf{n}dS$.

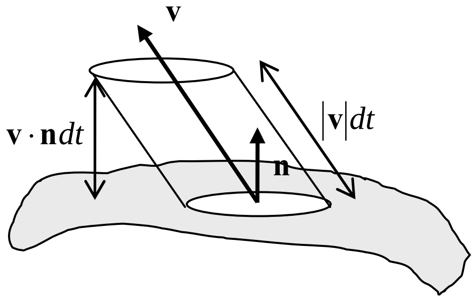


Figure 1.7.7: flow through a surface element

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

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

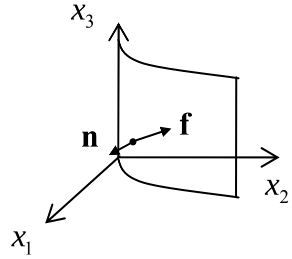


Figure 1.7.8: flux through a parabolic cylinder

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

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

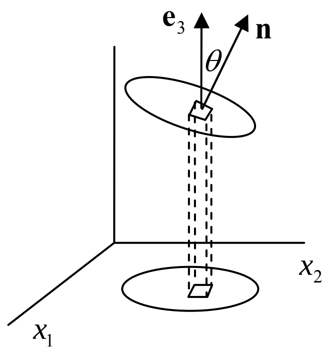


Figure 1.7.9: projection of a surface element onto a plane region

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)

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

where the Jacobian is now

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial t_1} & \frac{\partial x_2}{\partial t_1} & \frac{\partial x_3}{\partial t_1} \\ \frac{\partial x_1}{\partial t_2} & \frac{\partial x_2}{\partial t_2} & \frac{\partial x_3}{\partial t_2} \\ \frac{\partial x_1}{\partial t_3} & \frac{\partial x_2}{\partial t_3} & \frac{\partial x_3}{\partial t_3} \end{vmatrix} \quad \text{or} \quad J = \begin{vmatrix} \frac{\partial x_1}{\partial t_1} & \frac{\partial x_1}{\partial t_2} & \frac{\partial x_1}{\partial t_3} \\ \frac{\partial x_2}{\partial t_1} & \frac{\partial x_2}{\partial t_2} & \frac{\partial x_2}{\partial t_3} \\ \frac{\partial x_3}{\partial t_1} & \frac{\partial x_3}{\partial t_2} & \frac{\partial x_3}{\partial t_3} \end{vmatrix} \qquad (1.7.11)$$

so that

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

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)

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

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

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

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

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

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

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

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**)

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

Note that the term $\mathbf{n}\cdot\text{grad}\phi$ 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**)

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

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 $\mathrm{curl}\mathbf{f} = \mathbf{o}$.
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 $\mathrm{div}\mathbf{v} = 0$.

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]

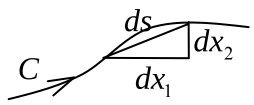


Figure 1.7.10: projection of a surface element onto a plane region

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

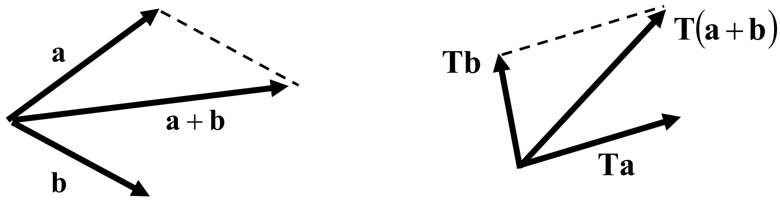


Figure 1.8.1: Linearity of the second order tensor

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

$$\mathbf{S}\mathbf{v} = \mathbf{T}\mathbf{v}$$

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.

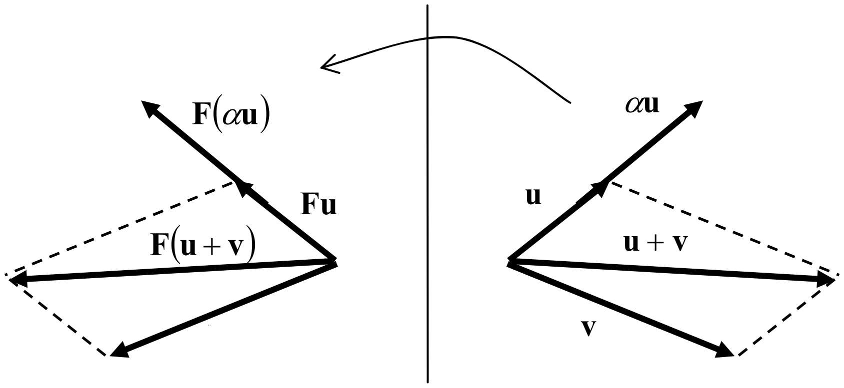


Figure 1.8.2: Mirror-imaging of vectors as a second order tensor mapping

■

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 $\mathbf{u}\times$

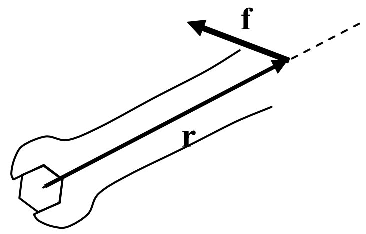


Figure 1.8.3: the force on a spanner

■

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

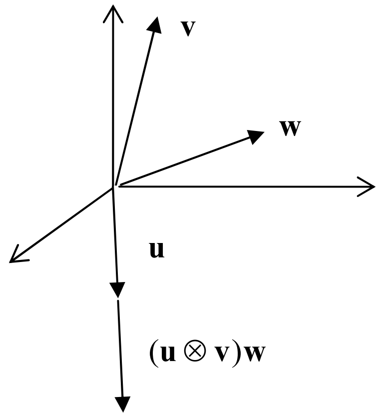


Figure 1.8.4: the dyad transformation

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

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

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

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

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**.

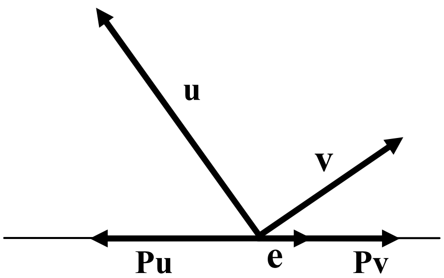


Figure 1.8.5: the projection tensor

■

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

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

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.

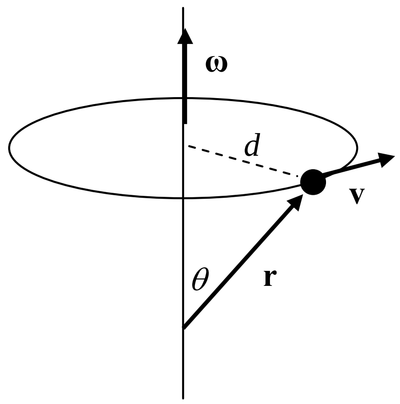


Figure 1.8.6: a particle in motion about an axis

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.

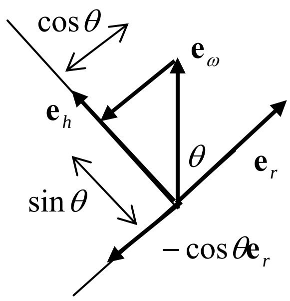


Figure 1.8.7: geometry of unit vectors for angular momentum calculation

■

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

$$\boxed{\begin{aligned} \mathbf{T} = T_{11}\mathbf{e}_1\otimes\mathbf{e}_1 + T_{12}\mathbf{e}_1\otimes\mathbf{e}_2 + T_{13}\mathbf{e}_1\otimes\mathbf{e}_3 \\ + T_{21}\mathbf{e}_2\otimes\mathbf{e}_1 + T_{22}\mathbf{e}_2\otimes\mathbf{e}_2 + T_{23}\mathbf{e}_2\otimes\mathbf{e}_3 \\ + T_{31}\mathbf{e}_3\otimes\mathbf{e}_1 + T_{32}\mathbf{e}_3\otimes\mathbf{e}_2 + T_{33}\mathbf{e}_3\otimes\mathbf{e}_3 \end{aligned}}$$

Second-order Cartesian Tensor (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}

$$\boxed{T_{ij} = \mathbf{e}_i\mathbf{T}\mathbf{e}_j}$$

Components of a Tensor (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 $\mathbf{n}$ 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

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

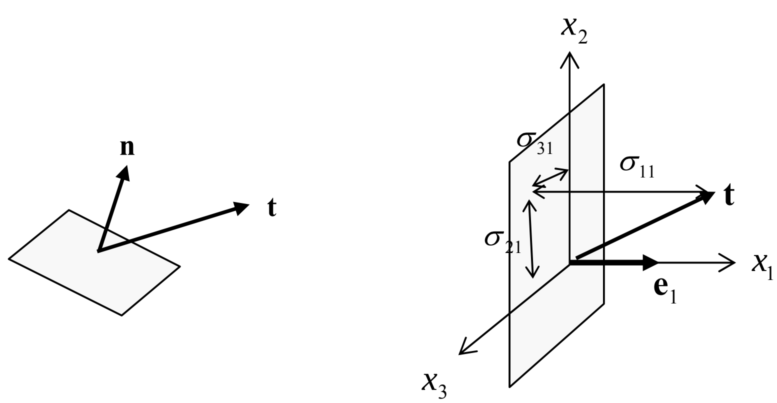


Figure 1.9.1: stress acting on a plane

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")
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	…	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

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

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 $\mathbf{A}$ 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

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

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}

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

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,

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

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

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

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,

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

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 $\mathbf{u}\otimes\mathbf{v}$,

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

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}

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

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

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

In summary then,

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

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, $\mathbf{T} \cdot\cdot \mathbf{S}$, $\boldsymbol{\mathsf{A}} \cdot\cdot \mathbf{b}$, 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,

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

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**,

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

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

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A} \qquad (1.10.17)$$

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:

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

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

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

1.10.7 Orthogonal Tensors

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

$$\mathbf{Qu} \cdot \mathbf{Qv} = \mathbf{u} \cdot \mathbf{v} \qquad (1.10.20)$$

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.

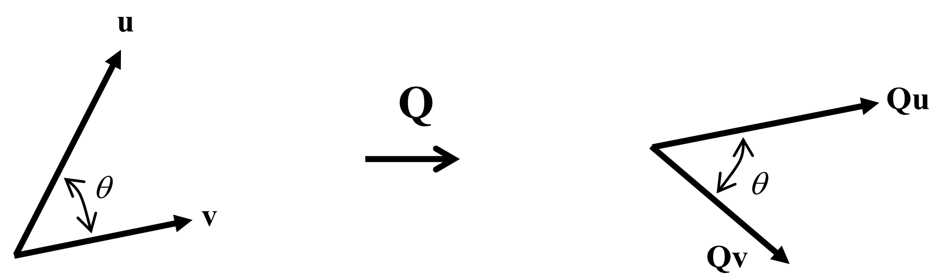


Figure 1.10.1: An orthogonal tensor

Since

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

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).

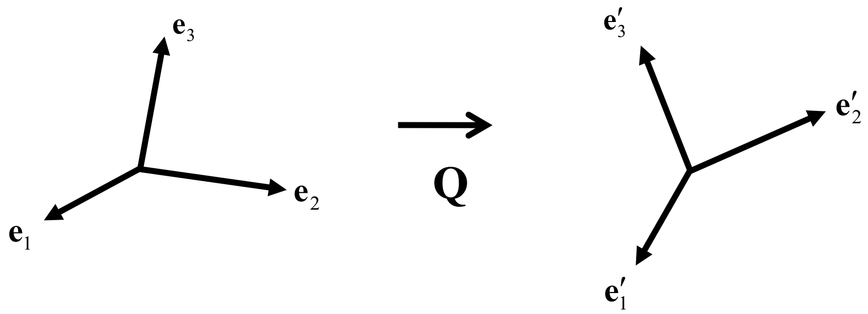


Figure 1.10.2: Rotation of a set of base vectors

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 **Q** transforms **v** into a second vector $\mathbf{v}'$, but this new vector has the *same components* with respect to the basis $\mathbf{e}'_i$, as **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, **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)

$$\mathbf{Q}\mathbf{v} = \begin{bmatrix} -1 \\ +1 \end{bmatrix}\mathbf{e}_i = \begin{bmatrix} +1 \\ +1 \end{bmatrix}\mathbf{e}'_i$$

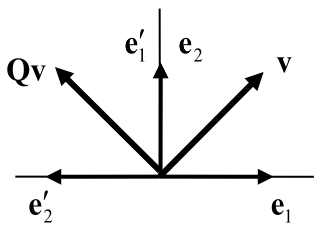


Figure 1.10.3: a rotated vector

■

Similarly, for a second order tensor **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}:

$$\begin{aligned} &\mathbf{S} : \mathbf{B} = \mathbf{S} : \mathbf{B}^{\mathrm{T}} = \mathbf{S} : \tfrac{1}{2}\left(\mathbf{B} + \mathbf{B}^{\mathrm{T}}\right) \\ &\mathbf{W} : \mathbf{B} = -\mathbf{W} : \mathbf{B}^{\mathrm{T}} = \mathbf{W} : \tfrac{1}{2}\left(-\mathbf{B}^{\mathrm{T}}\right) \\ &\mathbf{S} : \mathbf{W} = 0 \\ &\mathrm{tr}(\mathbf{SW}) = 0 \\ &\mathbf{v} \cdot \mathbf{Wv} = 0 \\ &\det \mathbf{W} = 0 \quad (\text{has no inverse}) \end{aligned} \tag{1.10.31}$$

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

$$\mathbf{W}\mathbf{u} = \boldsymbol{\omega} \times \mathbf{u} \tag{1.10.32}$$

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 $\mathbf{W}\mathbf{a} = \boldsymbol{\omega} \times \mathbf{a}$ 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

$$\mathbf{A} : \mathbf{v} \otimes \mathbf{v} > 0 \qquad (1.10.40)$$

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.

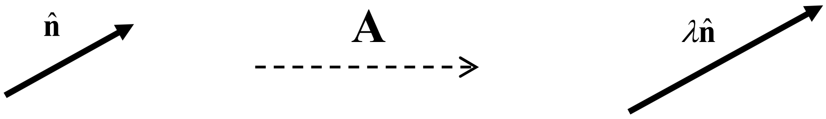


Figure 1.11.1: the action of a tensor A on a unit vector

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,

$$\begin{aligned}
&A_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j)\hat{n}_k\mathbf{e}_k - \lambda\delta_{pq}(\mathbf{e}_p \otimes \mathbf{e}_q)\hat{n}_r\mathbf{e}_r = 0 \\
&\to A_{ij}\hat{n}_j\mathbf{e}_i - \lambda\hat{n}_r\mathbf{e}_r = 0 \\
&\to (A_{ij}\hat{n}_j - \lambda\hat{n}_i)\mathbf{e}_i = 0
\end{aligned}$$

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

$$\det(\mathbf{A}-\lambda\mathbf{I}) = \det\begin{bmatrix} A_{11}-\lambda & A_{12} & A_{13} \\ A_{21} & A_{22}-\lambda & A_{23} \\ A_{31} & A_{32} & A_{33}-\lambda \end{bmatrix} = 0 \qquad (1.11.4)$$

Evaluating the determinant, one has the following cubic **characteristic equation** of **A**,

$$\boxed{\lambda^3 - \mathrm{I}_{\mathbf{A}}\lambda^2 + \mathrm{II}_{\mathbf{A}}\lambda - \mathrm{III}_{\mathbf{A}} = 0} \quad \textbf{Tensor Characteristic Equation} \quad (1.11.5)$$

where

$$\begin{aligned} \mathrm{I}_{\mathbf{A}} &= A_{ii} \\ &= \mathrm{tr}\mathbf{A} \\ \mathrm{II}_{\mathbf{A}} &= \tfrac{1}{2}\left(A_{ii}A_{jj} - A_{ji}A_{ij}\right) \\ &= \tfrac{1}{2}\left[(\mathrm{tr}\mathbf{A})^2 - \mathrm{tr}(\mathbf{A}^2)\right] \\ \mathrm{III}_{\mathbf{A}} &= \varepsilon_{ijk}A_{1i}A_{2j}A_{3k} \\ &= \det\mathbf{A} \end{aligned} \qquad (1.11.6)$$

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

$$\begin{aligned} \mathrm{I}_{\mathbf{A}} &= \lambda_1 + \lambda_2 + \lambda_3 \\ \mathrm{II}_{\mathbf{A}} &= \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 \\ \mathrm{III}_{\mathbf{A}} &= \lambda_1\lambda_2\lambda_3 \end{aligned} \qquad (1.11.7)$$

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

$$\begin{aligned} (A_{11}-\lambda)\hat{n}_1 + A_{12}\hat{n}_2 + A_{13}\hat{n}_3 &= 0 \\ A_{21}\hat{n}_1 + (A_{22}-\lambda)\hat{n}_2 + A_{23}\hat{n}_3 &= 0 \\ A_{31}\hat{n}_1 + A_{32}\hat{n}_2 + (A_{33}-\lambda)\hat{n}_3 &= 0 \end{aligned} \qquad (1.11.8)$$

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°.

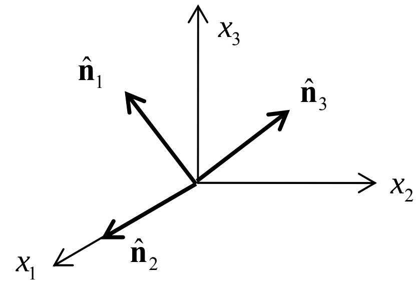


Figure 1.11.2: eigenvectors of the tensor T

■

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)

$$\mathbf{A} = A_{ij}\left(\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_j\right) \tag{1.11.9}$$

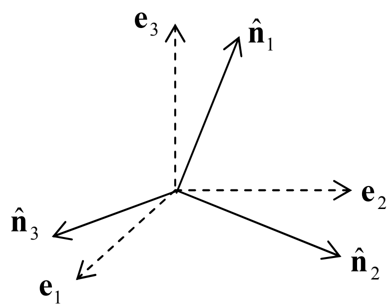


Figure 1.11.3: eigenvectors forming an orthonormal set

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:

$$\mathbf{A}^3 - \mathrm{I}_{\mathbf{A}} \mathbf{A}^2 + \mathrm{II}_{\mathbf{A}} \mathbf{A} - \mathrm{III}_{\mathbf{A}} \mathbf{I} = \mathbf{0} \qquad (1.11.15)$$

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:

$$\begin{aligned} & \mathbf{A}^3 : \mathbf{I} - \mathrm{I}_{\mathbf{A}} \mathbf{A}^2 : \mathbf{I} + \mathrm{II}_{\mathbf{A}} \mathbf{A} : \mathbf{I} - \mathrm{III}_{\mathbf{A}} \mathbf{I} : \mathbf{I} = 0 \\ \rightarrow \quad & \mathrm{tr}\, \mathbf{A}^3 - \mathrm{I}_{\mathbf{A}} \mathrm{tr}\, \mathbf{A}^2 + \mathrm{II}_{\mathbf{A}} \mathrm{tr}\, \mathbf{A} - 3\mathrm{III}_{\mathbf{A}} = 0 \\ \rightarrow \quad & \mathrm{tr}\, \mathbf{A}^3 - \mathrm{tr}\, \mathbf{A} \mathrm{tr}\, \mathbf{A}^2 + \tfrac{1}{2} \left[(\mathrm{tr}\, \mathbf{A})^2 - \mathrm{tr}\, \mathbf{A}^2 \right] \mathrm{tr}\, \mathbf{A} - 3\mathrm{III}_{\mathbf{A}} = 0 \\ \rightarrow \quad & \mathrm{III}_{\mathbf{A}} = \tfrac{1}{3} \left[\mathrm{tr}\, \mathbf{A}^3 - \tfrac{3}{2} \mathrm{tr}\, \mathbf{A}\, \mathrm{tr}\, \mathbf{A}^2 + \tfrac{1}{2} (\mathrm{tr}\, \mathbf{A})^3 \right] \end{aligned} \qquad (1.11.16)$$

The three invariants of a tensor can now be listed as

$$\boxed{\begin{aligned} \mathrm{I}_{\mathbf{A}} &= \mathrm{tr} \mathbf{A} \\ \mathrm{II}_{\mathbf{A}} &= \tfrac{1}{2} \left[(\mathrm{tr} \mathbf{A})^2 - \mathrm{tr}(\mathbf{A}^2) \right] \\ \mathrm{III}_{\mathbf{A}} &= \tfrac{1}{3} \left[\mathrm{tr}\, \mathbf{A}^3 - \tfrac{3}{2} \mathrm{tr}\, \mathbf{A}\, \mathrm{tr}\, \mathbf{A}^2 + \tfrac{1}{2} (\mathrm{tr}\, \mathbf{A})^3 \right] \end{aligned}} \qquad \textbf{Invariants of a Tensor} \qquad (1.11.17)$$

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

$$s^3 - J_1 s^2 - J_2 s - J_3 = 0 \tag{1.11.18}$$

and the deviatoric invariants are[3]

$$\begin{aligned} J_1 &= \mathrm{tr}(\mathrm{dev}\mathbf{A}) = s_1 + s_2 + s_3 \\ J_2 &= -\tfrac{1}{2}\left[(\mathrm{tr}(\mathrm{dev}\mathbf{A}))^2 - \mathrm{tr}\left((\mathrm{dev}\mathbf{A})^2\right)\right] = -(s_1 s_2 + s_2 s_3 + s_3 s_1) \\ J_3 &= \det(\mathrm{dev}\mathbf{A}) = s_1 s_2 s_3 \end{aligned} \tag{1.11.19}$$

From Eqn. 1.10.37,

$$J_1 = 0 \tag{1.11.20}$$

The second invariant can also be expressed in the useful forms {▲Problem 4}

$$J_2 = \tfrac{1}{2}\left(s_1^2 + s_2^2 + s_3^2\right), \tag{1.11.21}$$

and, in terms of the eigenvalues of **A**, {▲Problem 5}

$$J_2 = \frac{1}{6}\left[(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2\right]. \tag{1.11.22}$$

Further, the deviatoric invariants are related to the tensor invariants through {▲Problem 6}

$$J_2 = \tfrac{1}{3}\left(\mathrm{I}_\mathbf{A}^2 - 3\mathrm{II}_\mathbf{A}\right), \quad J_3 = \tfrac{1}{27}\left(2\mathrm{I}_\mathbf{A}^3 - 9\mathrm{I}_\mathbf{A}\mathrm{II}_\mathbf{A} + 27\mathrm{III}_\mathbf{A}\right) \tag{1.11.23}$$

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 **T** repeatedly gives $\mathbf{T}^n\hat{\mathbf{n}} = \lambda^n\hat{\mathbf{n}}$. It follows that, if **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 **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

$$\mathbf{T}^{1/2} = \sum_{i=1}^{3} \sqrt{\lambda_i}\,\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \tag{1.11.24}$$

and the inverse is

$$\mathbf{T}^{-1} = \sum_{i=1}^{3} (1/\lambda_i)\hat{\mathbf{n}}_i \otimes \hat{\mathbf{n}}_i \tag{1.11.25}$$

These new tensors are also positive definite.

1.11.7 Polar Decomposition of Tensors

Any (non-singular second-order) tensor **F** can be split up multiplicatively into an arbitrary proper orthogonal tensor **R** ($\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$, $\det\mathbf{R} = 1$) and a tensor **U** as follows:

$$\boxed{\mathbf{F} = \mathbf{R}\mathbf{U}} \qquad \textbf{Polar Decomposition} \tag{1.11.26}$$

The consequence of this is that any transformation of a vector **a** according to $\mathbf{Fa}$ can be decomposed into two transformations, one involving a transformation **U**, followed by a rotation **R**.

The decomposition is not, in general, unique; one can often find more than one orthogonal tensor **R** which will satisfy the above relation. In practice, **R** is chosen such that **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 **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}:

$$\begin{aligned} \mathbf{I}\otimes\mathbf{I}:\mathbf{A} &= (\mathrm{tr}\mathbf{A})\mathbf{I} \\ \mathbf{I}\otimes\mathbf{I}:\mathbf{A}:\mathbf{A} &= (\mathrm{tr}\mathbf{A})^2 \\ \mathsf{I}:\mathbf{A}:\mathbf{A} &= \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\mathbf{A}) \\ \bar{\mathsf{I}}:\mathbf{A}:\mathbf{A} &= \mathrm{tr}\mathbf{A}^2 \end{aligned} \qquad (1.12.6)$$

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 $\mathbf{E} = \varepsilon_{ijk}(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)$ 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)

$$\boxed{T_{ij} = Q_{ip}Q_{jq}T'_{pq}, \quad T'_{ij} = Q_{pi}Q_{qj}T_{pq}} \quad \textbf{Tensor Transformation Formulae} \quad (1.13.5)$$

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 $\mathbf{S}$ 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 $\mathbf{S}$ 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 $\mathbf{S}$ symmetric, $S_{12} = S_{21}$, and this simplifies to

$$\boxed{\begin{aligned} S'_{11} &= S_{11}\cos^2\theta + S_{22}\sin^2\theta + S_{12}\sin 2\theta \\ S'_{22} &= S_{11}\sin^2\theta + S_{22}\cos^2\theta - S_{12}\sin 2\theta \\ S'_{12} &= (S_{22} - S_{11})\sin\theta\cos\theta + S_{12}\cos 2\theta \end{aligned}} \quad \textbf{The Mohr Transformation} \quad (1.13.7)$$

■

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,

$$\mathsf{I},\quad \bar{\mathsf{I}},\quad \mathbf{I}\otimes\mathbf{I}$$

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:

$$\boxed{\mathsf{C} = \lambda\mathbf{I}\otimes\mathbf{I} + \mu\mathsf{I} + \gamma\bar{\mathsf{I}}} \quad \textbf{Most General Isotropic Fourth-Order Tensor} \qquad (1.13.11)$$

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

$$\frac{d\mathbf{T}}{dt} = \frac{dT_{ij}}{dt} \mathbf{e}_i \otimes \mathbf{e}_j \qquad (1.14.1)$$

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

$$\begin{aligned}
\frac{d}{dt}(\mathbf{T} + \mathbf{B}) &= \frac{d\mathbf{T}}{dt} + \frac{d\mathbf{B}}{dt} \\
\frac{d}{dt}(\alpha(t)\mathbf{T}) &= \alpha \frac{d\mathbf{T}}{dt} + \frac{d\alpha}{dt}\mathbf{T} \\
\frac{d}{dt}(\mathbf{T}\mathbf{a}) &= \mathbf{T}\frac{d\mathbf{a}}{dt} + \frac{d\mathbf{T}}{dt}\mathbf{a} \\
\frac{d}{dt}(\mathbf{T}\mathbf{B}) &= \mathbf{T}\frac{d\mathbf{B}}{dt} + \frac{d\mathbf{T}}{dt}\mathbf{B} \\
\frac{d}{dt}(\mathbf{T}^{\mathrm{T}}) &= \left(\frac{d\mathbf{T}}{dt}\right)^{\mathrm{T}}
\end{aligned}$$

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

$$\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}} = -\mathbf{Q}\dot{\mathbf{Q}}^{\mathrm{T}} = -\left(\dot{\mathbf{Q}}\mathbf{Q}^{\mathrm{T}}\right)^{\mathrm{T}} \qquad (1.14.2)$$

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

$$\boxed{\operatorname{grad}\mathbf{a} \equiv \frac{\partial \mathbf{a}}{\partial x_j} \otimes \mathbf{e}_j = \frac{\partial a_i}{\partial x_j}\mathbf{e}_i \otimes \mathbf{e}_j} \qquad \textbf{Gradient of a Vector Field} \qquad (1.14.3)$$

In matrix notation,

$$\operatorname{grad}\mathbf{a} = \begin{bmatrix} \dfrac{\partial a_1}{\partial x_1} & \dfrac{\partial a_1}{\partial x_2} & \dfrac{\partial a_1}{\partial x_3} \\ \dfrac{\partial a_2}{\partial x_1} & \dfrac{\partial a_2}{\partial x_2} & \dfrac{\partial a_2}{\partial x_3} \\ \dfrac{\partial a_3}{\partial x_1} & \dfrac{\partial a_3}{\partial x_2} & \dfrac{\partial a_3}{\partial x_3} \end{bmatrix} \qquad (1.14.4)$$

One then has

$$\begin{aligned} \operatorname{grad}\mathbf{a}\, d\mathbf{x} &= \frac{\partial a_i}{\partial x_j}\mathbf{e}_i \otimes \mathbf{e}_j \left(dx_k \mathbf{e}_k\right) \\ &= \frac{\partial a_i}{\partial x_j} dx_j \mathbf{e}_i \\ &= d\mathbf{a} \\ &= \mathbf{a}(\mathbf{x} + d\mathbf{x}) - \mathbf{a}(d\mathbf{x}) \end{aligned} \qquad (1.14.5)$$

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

$$\operatorname{grad}\mathbf{a}\,\mathbf{e} = \left(\frac{d\mathbf{a}}{dx}\right)_{\text{in } \mathbf{e} \text{ direction}} \qquad (1.14.6)$$

Thus the gradient of a vector field **a** is a second-order tensor which transforms a unit vector into a vector describing the gradient of **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

$$\frac{d\mathbf{a}}{ds} = \frac{\partial \mathbf{a}}{\partial x_j}\frac{dx_j}{ds} = \frac{\partial \mathbf{a}}{\partial x_j}\left(\boldsymbol{\tau}\cdot\mathbf{e}_j\right) = \left(\frac{\partial \mathbf{a}}{\partial x_j}\otimes\mathbf{e}_j\right)\boldsymbol{\tau} = \mathrm{grad}\,\mathbf{a}\,\boldsymbol{\tau}\,.$$

Although for a scalar field $\mathrm{grad}\phi$ 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,

$$\left(\nabla\otimes\mathbf{a}\right)^{\mathrm{T}} = \mathrm{grad}\mathbf{a} \tag{1.14.7}$$

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 $\mathrm{grad}\mathbf{a}$ (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 $\nabla\otimes\mathbf{a} = \mathrm{grad}\mathbf{a} = \left(\partial a_i/\partial x_j\right)\mathbf{e}_i\otimes\mathbf{e}_j$.

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

$$\begin{aligned} d\mathbf{x} &= d\mathbf{X} + \mathbf{u}(\mathbf{X}+d\mathbf{X}) - \mathbf{u}(\mathbf{X}) \\ &= d\mathbf{X} + d\mathbf{u}(\mathbf{X}) \\ &= d\mathbf{X} + \mathrm{grad}\mathbf{u}\,d\mathbf{X} \end{aligned}$$

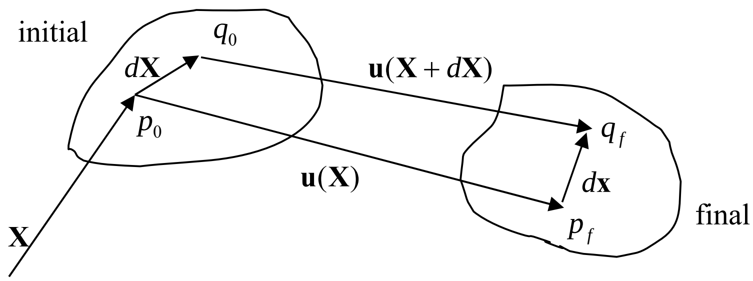


Figure 1.14.1: displacement of material particles

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

$$\operatorname{grad}\mathbf{u} = \frac{\partial u_i}{\partial X_j} = \begin{bmatrix} 0 & 2kX_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 2kX_2 \mathbf{e}_1 \otimes \mathbf{e}_2$$

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 $d\mathbf{x} = 2k\alpha\, d\alpha\mathbf{e}_1 + \alpha\mathbf{e}_2$.

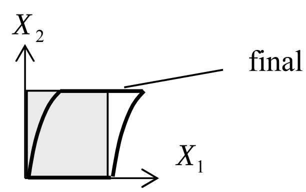


Figure 1.14.2: deformation of a box

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},

$$\boxed{\operatorname{div}\mathbf{a} = \operatorname{tr}(\operatorname{grad}\mathbf{a}) = \operatorname{grad}\mathbf{a} : \mathbf{I} = \nabla \cdot \mathbf{a}}$$ **Divergence of a Vector Field** (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

$$\boxed{\operatorname{grad}\mathbf{T} = \frac{\partial \mathbf{T}}{\partial x_k} \otimes \mathbf{e}_k = \frac{\partial T_{ij}}{\partial x_k} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k}$$ **Gradient of a Tensor Field** (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

$$\boxed{\begin{aligned} \operatorname{div}\mathbf{T} = \operatorname{grad}\mathbf{T} : \mathbf{I} &= \frac{\partial \mathbf{T}}{\partial x_i} \mathbf{e}_i = \frac{\partial (T_{jk} \mathbf{e}_j \otimes \mathbf{e}_k)}{\partial x_i} \mathbf{e}_i \\ &= \frac{\partial T_{ij}}{\partial x_j} \mathbf{e}_i \end{aligned}}$$ **Divergence of a Tensor** (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}:

$$\begin{aligned}
\text{grad}(\phi\mathbf{v}) &= \phi\,\text{grad}\mathbf{v} + \mathbf{v} \otimes \text{grad}\phi \\
\text{grad}(\mathbf{u} \cdot \mathbf{v}) &= (\text{grad}\mathbf{u})^{\mathrm{T}}\mathbf{v} + (\text{grad}\mathbf{v})^{\mathrm{T}}\mathbf{u} \\
\text{div}(\mathbf{u} \otimes \mathbf{v}) &= (\text{grad}\,\mathbf{u})\mathbf{v} + (\text{div}\,\mathbf{v})\mathbf{u} \\
\text{curl}(\mathbf{u} \times \mathbf{v}) &= \mathbf{u}\,\text{div}\mathbf{v} - \mathbf{v}\,\text{div}\mathbf{u} + (\text{grad}\mathbf{u})\mathbf{v} - (\text{grad}\mathbf{v})\mathbf{u}
\end{aligned} \qquad (1.14.15)$$

$$\begin{aligned}
\text{div}(\phi\mathbf{A}) &= \mathbf{A}\,\text{grad}\phi + \phi\,\text{div}\mathbf{A} \\
\text{div}(\mathbf{A}\mathbf{v}) &= \mathbf{v} \cdot \text{div}\mathbf{A}^{\mathrm{T}} + \text{tr}(\mathbf{A}\,\text{grad}\mathbf{v}) \\
\text{div}(\mathbf{A}\mathbf{B}) &= \mathbf{A}\,\text{div}\mathbf{B} + \text{grad}\mathbf{A} : \mathbf{B} \\
\text{div}(\mathbf{A}(\phi\mathbf{B})) &= \phi\,\text{div}(\mathbf{A}\mathbf{B}) + \mathbf{A}(\mathbf{B}\,\text{grad}\phi) \\
\text{grad}(\phi\mathbf{A}) &= \phi\,\text{grad}\mathbf{A} + \mathbf{A} \otimes \text{grad}\phi
\end{aligned} \qquad (1.11.16)$$

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 $\text{grad}\mathbf{u} = (\nabla \otimes \mathbf{u})^{\mathrm{T}}$ with

$$\begin{aligned}
\text{grad}\mathbf{u} &= \left[\left(\mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \mathbf{e}_z \frac{\partial}{\partial z}\right) \otimes \left(u_r \mathbf{e}_r + u_\theta \mathbf{e}_\theta + u_z \mathbf{e}_z\right)\right]^{\mathrm{T}} \\
&= \frac{\partial u_r}{\partial r}\mathbf{e}_r \otimes \mathbf{e}_r + \left(\frac{1}{r}\frac{\partial u_r}{\partial \theta} - \frac{u_\theta}{r}\right)\mathbf{e}_r \otimes \mathbf{e}_\theta + \frac{\partial u_r}{\partial z}\mathbf{e}_r \otimes \mathbf{e}_z \\
&\quad + \frac{\partial u_\theta}{\partial r}\mathbf{e}_\theta \otimes \mathbf{e}_r + \left(\frac{1}{r}\frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r}\right)\mathbf{e}_\theta \otimes \mathbf{e}_\theta + \frac{\partial u_\theta}{\partial z}\mathbf{e}_\theta \otimes \mathbf{e}_z \\
&\quad + \frac{\partial u_z}{\partial r}\mathbf{e}_z \otimes \mathbf{e}_r + \frac{1}{r}\frac{\partial u_z}{\partial \theta}\mathbf{e}_z \otimes \mathbf{e}_\theta + \frac{\partial u_z}{\partial z}\mathbf{e}_z \otimes \mathbf{e}_z
\end{aligned} \qquad (1.14.18)$$

and from 1.6.30, 1.14.12, the divergence of a tensor in cylindrical coordinates is {▲Problem 6}

$$\begin{aligned}
\text{div}\mathbf{A} = \nabla \cdot \mathbf{A}^{\mathrm{T}} &= \left(\frac{\partial A_{rr}}{\partial r} + \frac{1}{r}\frac{\partial A_{r\theta}}{\partial \theta} + \frac{\partial A_{rz}}{\partial z} + \frac{A_{rr} - A_{\theta\theta}}{r}\right)\mathbf{e}_r \\
&\quad + \left(\frac{\partial A_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial A_{\theta\theta}}{\partial \theta} + \frac{\partial A_{\theta z}}{\partial z} + \frac{A_{\theta r} + A_{r\theta}}{r}\right)\mathbf{e}_\theta \\
&\quad + \left(\frac{\partial A_{zr}}{\partial r} + \frac{A_{zr}}{r} + \frac{1}{r}\frac{\partial A_{z\theta}}{\partial \theta} + \frac{\partial A_{zz}}{\partial z}\right)\mathbf{e}_z
\end{aligned} \qquad (1.14.19)$$

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

$$\int_S T_{ij\cdots k} n_k \, dS = \int_V \frac{\partial T_{ij\cdots k}}{\partial x_k} dV \qquad (1.14.20)$$

For a second order tensor,

$$\int_S \mathbf{Tn}\,dS = \int_V \operatorname{div}\mathbf{T}\,dV, \qquad \int_S T_{ij} n_j\,dS = \int_V \frac{\partial T_{ij}}{\partial x_j}\,dV \tag{1.14.21}$$

One then has the important identities {▲Problem 7}

$$\begin{aligned}\int_S (\phi\mathbf{T})\mathbf{n}\,dS &= \int_V \operatorname{div}(\phi\mathbf{T})\,dV \\ \int_S \mathbf{u}\otimes\mathbf{n}\,dS &= \int_V \operatorname{grad}\mathbf{u}\,dV \\ \int_S \mathbf{u}\cdot\mathbf{Tn}\,dS &= \int_V \operatorname{div}(\mathbf{T}^{\mathrm{T}}\mathbf{u})\,dV\end{aligned} \tag{1.14.22}$$

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,

$$\nabla \mathbf{v}(\mathbf{x}) = \frac{\partial v_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j \tag{1.14.26}$$

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 $(\text{grad}\mathbf{v})\mathbf{v}$.
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 $\text{grad}\mathbf{u}$, sketch a few (undeformed) line elements of material and their positions in the deformed configuration.
4. Use the matrix form of $\text{grad}\mathbf{u}$ and $\nabla \otimes \mathbf{u}$ to show that the definitions
   (i) $\text{div}\mathbf{a} = \text{tr}(\text{grad}\mathbf{a})$
   (ii) $\text{curl}\mathbf{a} = 2\boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is the axial vector of the skew part of $\text{grad}\mathbf{a}$
   agree with the definitions 1.6.17, 1.6.21 given for Cartesian coordinates.
5. Prove the following:
   (i) $\text{grad}(\phi\mathbf{v}) = \phi\text{grad}\mathbf{v} + \mathbf{v} \otimes \text{grad}\phi$
   (ii) $\text{grad}(\mathbf{u} \cdot \mathbf{v}) = (\text{grad}\mathbf{u})^{\mathrm{T}} \mathbf{v} + (\text{grad}\mathbf{v})^{\mathrm{T}} \mathbf{u}$
   (iii) $\text{div}(\mathbf{u} \otimes \mathbf{v}) = (\text{grad}\,\mathbf{u})\mathbf{v} + (\text{div}\,\mathbf{v})\mathbf{u}$
   (iv) $\text{curl}(\mathbf{u} \times \mathbf{v}) = \mathbf{u}\text{div}\mathbf{v} - \mathbf{v}\text{div}\mathbf{u} + (\text{grad}\mathbf{u})\mathbf{v} - (\text{grad}\mathbf{v})\mathbf{u}$
   (v) $\text{div}(\phi\mathbf{A}) = \mathbf{A}\text{grad}\phi + \phi\text{div}\mathbf{A}$
   (vi) $\text{div}(\mathbf{A}\mathbf{v}) = \mathbf{v} \cdot \text{div}\mathbf{A}^{\mathrm{T}} + \text{tr}(\mathbf{A}\text{grad}\mathbf{v})$
   (vii) $\text{div}(\mathbf{A}\mathbf{B}) = \mathbf{A}\text{div}\mathbf{B} + \text{grad}\mathbf{A} : \mathbf{B}$
   (viii) $\text{div}(\mathbf{A}(\phi\mathbf{B})) = \phi\text{div}(\mathbf{A}\mathbf{B}) + \mathbf{A}(\mathbf{B}\text{grad}\phi)$
   (ix) $\text{grad}(\phi\mathbf{A}) = \phi\text{grad}\mathbf{A} + \mathbf{A} \otimes \text{grad}\phi$
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 **T** is defined to be a second-order tensor with these partial derivatives as its components:

$$\boxed{\frac{\partial \phi}{\partial \mathbf{T}} \equiv \frac{\partial \phi}{\partial T_{ij}} \mathbf{e}_i \otimes \mathbf{e}_j} \qquad \textbf{Partial Derivative with respect to a Tensor} \qquad (1.15.3)$$

The quantity $\partial \phi(\mathbf{T}) / \partial \mathbf{T}$ is also called the gradient of ϕ with respect to **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 **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}

$$\boxed{\begin{array}{ccc} \dfrac{\partial (\mathrm{tr}\mathbf{A})}{\partial \mathbf{A}} = \mathbf{I} & \dfrac{\partial (\mathrm{tr}\mathbf{A}^2)}{\partial \mathbf{A}} = 2\mathbf{A}^{\mathrm{T}} & \dfrac{\partial (\mathrm{tr}\mathbf{A}^3)}{\partial \mathbf{A}} = 3(\mathbf{A}^2)^{\mathrm{T}} \\ \dfrac{\partial ((\mathrm{tr}\mathbf{A})^2)}{\partial \mathbf{A}} = 2(\mathrm{tr}\mathbf{A})\mathbf{I} & \dfrac{\partial ((\mathrm{tr}\mathbf{A})^3)}{\partial \mathbf{A}} = 3(\mathrm{tr}\mathbf{A})^2 \mathbf{I} & \end{array}} \qquad (1.15.5)$$

Derivatives of Trace Functions

From these and 1.11.17, one can evaluate the derivatives of the invariants {▲Problem 2}:

$$\boxed{\begin{aligned} \frac{\partial \mathrm{I}_{\mathbf{A}}}{\partial \mathbf{A}} &= \mathbf{I} \\ \frac{\partial \mathrm{II}_{\mathbf{A}}}{\partial \mathbf{A}} &= \mathrm{I}_{\mathbf{A}} \mathbf{I} - \mathbf{A}^{\mathrm{T}} \\ \frac{\partial \mathrm{III}_{\mathbf{A}}}{\partial \mathbf{A}} &= (\mathbf{A}^{\mathrm{T}})^2 - \mathrm{I}_{\mathbf{A}} \mathbf{A}^{\mathrm{T}} + \mathrm{II}_{\mathbf{A}} \mathbf{I} = \mathrm{III}_{\mathbf{A}} \mathbf{A}^{-\mathrm{T}} \end{aligned}} \qquad \textbf{Derivatives of the Invariants} \qquad (1.15.6)$$

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

$$\begin{aligned}\overline{A}_{12} &= \frac{1}{2}(A_{12} + A_{21}) = A_{12} = A_{21}\\ \overline{A}_{13} &= \frac{1}{2}(A_{13} + A_{31}) = A_{13} = A_{31}\\ \overline{A}_{23} &= \frac{1}{2}(A_{23} + A_{32}) = A_{23} = A_{32}\end{aligned}$$

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

$$\begin{aligned}
& \frac{\partial A_{im}^{-1}}{\partial A_{kl}} A_{mj} = -A_{im}^{-1} \frac{\partial A_{mj}}{\partial A_{kl}} && (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l) \\
\rightarrow \quad & \frac{\partial A_{im}^{-1}}{\partial A_{kl}} A_{mj} A_{jn}^{-1} = -A_{im}^{-1} \frac{\partial A_{mj}}{\partial A_{kl}} A_{jn}^{-1} \\
\rightarrow \quad & \frac{\partial A_{im}^{-1}}{\partial A_{kl}} \delta_{mn} = -A_{im}^{-1} \delta_{mk} \delta_{jl} A_{jn}^{-1} \\
\rightarrow \quad & \frac{\partial A_{ij}^{-1}}{\partial A_{kl}} = -A_{ik}^{-1} A_{lj}^{-1} && (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l)
\end{aligned} \tag{1.15.18}$$

■

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 $\mathbf{B} = \varepsilon \mathbf{A}^{-1}\mathbf{T}$,

$$
\begin{aligned}
\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] &= \det \mathbf{A} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left(1+\lambda_1|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\lambda_2|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\lambda_3|_{\varepsilon\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left(1+\varepsilon\lambda_1|_{\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\varepsilon\lambda_2|_{\mathbf{A}^{-1}\mathbf{T}}\right)\left(1+\varepsilon\lambda_3|_{\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A} \left(\lambda_1|_{\mathbf{A}^{-1}\mathbf{T}} + \lambda_2|_{\mathbf{A}^{-1}\mathbf{T}} + \lambda_3|_{\mathbf{A}^{-1}\mathbf{T}}\right) \\
&= \det \mathbf{A}\ \mathrm{tr}\left(\mathbf{A}^{-1}\mathbf{T}\right)
\end{aligned}
$$

and, from (1.10.10e),

$$
\partial_{\mathbf{A}}(\det \mathbf{A})[\mathbf{T}] = \det \mathbf{A}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right) \qquad (1.15.24)
$$

■

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

$$
\begin{aligned}
\partial_{\mathbf{x}}\mathbf{f}(\mathbf{x})[\mathbf{u}] &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathbf{f}(\mathbf{z}(\varepsilon)) \qquad (\mathbf{z} = \mathbf{x} + \varepsilon\mathbf{u}) \\
&= \left(\frac{\partial \mathbf{f}(\mathbf{z})}{\partial \mathbf{z}} \frac{d\mathbf{z}}{d\varepsilon}\right)_{\varepsilon=0} \\
&= \mathbf{K}\mathbf{u}
\end{aligned}
\qquad (1.15.25)
$$

where $\mathbf{K}$, called the **tangent matrix** of the system, is

$$
\mathbf{K} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & & \partial f_2/\partial x_n \\ \vdots & & & \vdots \\ \partial f_n/\partial x_1 & & \cdots & \partial f_n/\partial x_n \end{bmatrix}, \qquad \partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}] = (\mathrm{grad}\mathbf{f})\mathbf{u}
$$

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 $\mathbf{T}$:

$$\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (\mathbf{A}+\varepsilon\mathbf{T})^{-1}$$

To evaluate this, note that $\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\mathbf{A}\right)[\mathbf{T}] = \partial_{\mathbf{A}}(\mathbf{I})[\mathbf{T}] = \mathbf{0}$, since **I** is independent of **A**. The product rule then gives $\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}]\mathbf{A} = -\mathbf{A}^{-1}\partial_{\mathbf{A}}(\mathbf{A})[\mathbf{T}]$, so that

$$\partial_{\mathbf{A}}\left(\mathbf{A}^{-1}\right)[\mathbf{T}] = -\mathbf{A}^{-1}\partial_{\mathbf{A}}\mathbf{A}[\mathbf{T}]\mathbf{A}^{-1} = -\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1} \qquad (1.15.26)$$

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)

$$\mathbf{f}(\mathbf{x}+\mathbf{u}) \approx \mathbf{f}(\mathbf{x}) + \partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}], \qquad (1.15.27)$$

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,

$$\boxed{\partial_{\mathbf{x}}\mathbf{f}[\mathbf{u}] = \partial_{\mathbf{B}}\hat{\mathbf{f}}(\mathbf{B})[\partial_{\mathbf{x}}\mathbf{B}[\mathbf{u}]]} \qquad \textbf{Chain Rule} \qquad (1.15.28)$$

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,

$$\begin{aligned}\partial_{\mathbf{A}}\left(\det\mathbf{A}^{-1}\right)[\mathbf{T}] &= \partial_{\mathbf{B}}(\det\mathbf{B})[\partial_{\mathbf{A}}\mathbf{A}^{-1}[\mathbf{T}]] \\ &= (\det\mathbf{B})\left(\mathbf{B}^{-\mathrm{T}} : \left(-\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}\right)\right) \\ &= -\det\mathbf{A}^{-1}\left(\mathbf{A}^{\mathrm{T}} : \left(\mathbf{A}^{-1}\mathbf{T}\mathbf{A}^{-1}\right)\right) \\ &= -\det\mathbf{A}^{-1}\left(\mathbf{A}^{-\mathrm{T}} : \mathbf{T}\right)\end{aligned} \qquad (1.15.29)$$

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 $\mathbf{C} = \mathbf{U}\mathbf{U}$ 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).

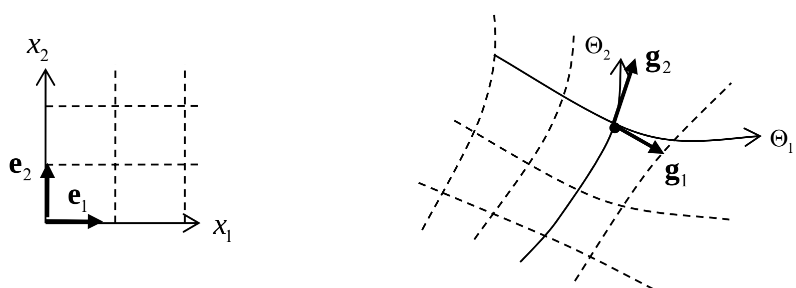


Figure 1.16.1: A Cartesian coordinate system and a curvilinear coordinate system

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.

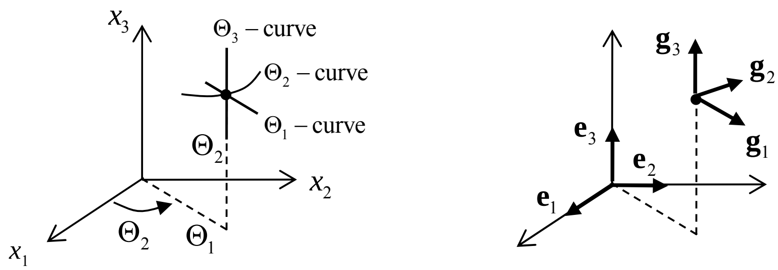


Figure 1.16.2: Cylindrical Coordinates

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.

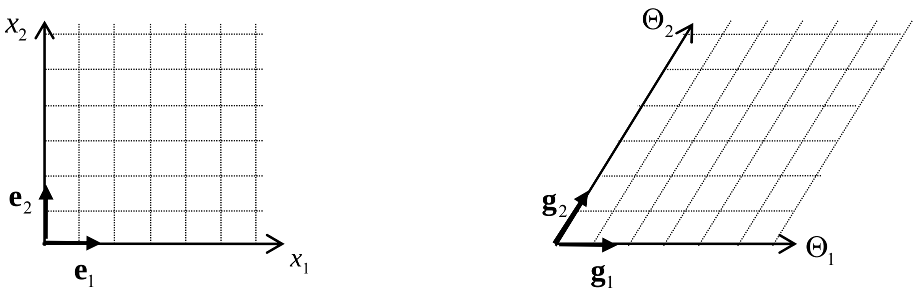


Figure 1.16.3: Deformation of a Square into a Parallelogram

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.

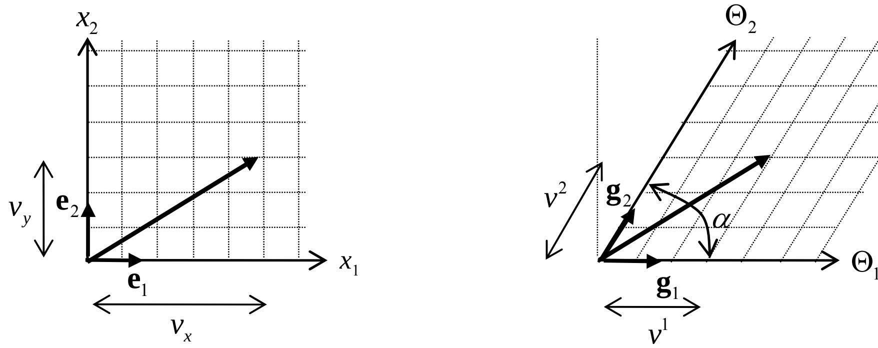


Figure 1.16.4: A Cartesian coordinate system and an oblique coordinate system

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 $\mathbf{g}_2$ $(\mathbf{g}^1\cdot\mathbf{g}_2 = 0)$, and the base vector $\mathbf{g}^2$ is perpendicular to $\mathbf{g}_1$ $(\mathbf{g}_1\cdot\mathbf{g}^2 = 0)$, 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 $|\mathbf{g}^1| = |\mathbf{g}^2| = 1/\sin\alpha$.

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

$$\begin{aligned} v_1 &= v_x \\ v_2 &= \cos\alpha \, v_x + \sin\alpha \, v_y \end{aligned} \tag{1.16.9}$$

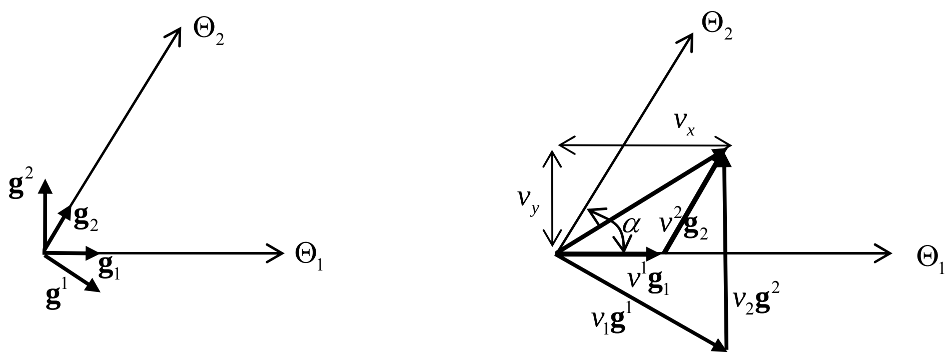


Figure 1.16.5: 2 sets of basis vectors; (a) covariant and contravariant base vectors, (b) covariant and contravariant components of a vector

Now one can evaluate the quantity

$$\begin{aligned} v_1 v^1 + v_2 v^2 &= \left(v_x\right)\left(v_x - \frac{1}{\tan\alpha} v_y\right) + \left(\cos\alpha \, v_x + \sin\alpha \, v_y\right)\left(\frac{1}{\sin\alpha} v_y\right) \\ &= v_x^2 + v_y^2 \end{aligned} \tag{1.16.10}$$

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)

$$\begin{aligned} \Theta^i &= \Theta^i\left(x^1, x^2, x^3\right) \\ x^i &= x^i(\Theta^1, \Theta^2, \Theta^3) \end{aligned} \qquad (1.16.15)$$

For example, the transformation equations for the oblique coordinate system of Fig. 1.16.4 are

$$\begin{aligned} &\Theta^1 = x^1 - \frac{1}{\tan\alpha} x^2, \quad \Theta^2 = \frac{1}{\sin\alpha} x^2, \quad \Theta^3 = x^3 \\ &x^1 = \Theta^1 + \cos\alpha\,\Theta^2, \quad x^2 = \sin\alpha\,\Theta^2, \quad x^3 = \Theta^3 \end{aligned} \qquad (1.16.16)$$

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.

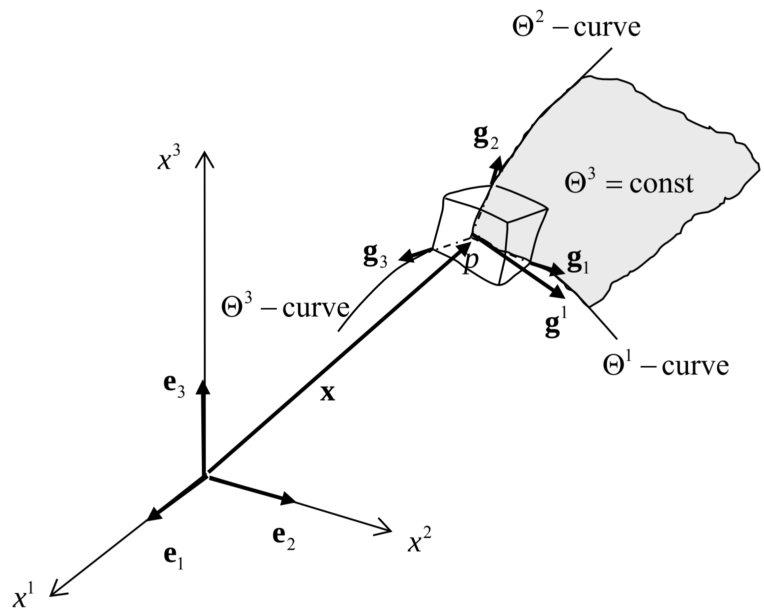


Figure 1.16.6: curvilinear coordinate system and coordinate curves

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_i^j 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}

$$\mathbf{g}_i = g_{ij}\mathbf{g}^j$$
$$\mathbf{g}^i = g^{ij}\mathbf{g}_j \qquad (1.16.29)$$

and {▲Problem 3}

$$g^{ij}g_{kj} = \delta^i_k \equiv g^i_k \qquad (1.16.30)$$

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:

$$g_{ij} = \mathbf{g}_i \cdot \mathbf{g}_j = \frac{\partial x^k}{\partial \Theta^i}\frac{\partial x^k}{\partial \Theta^j}, \qquad g^{ij} = \mathbf{g}^i \cdot \mathbf{g}^j = \frac{\partial \Theta^i}{\partial x^k}\frac{\partial \Theta^j}{\partial x^k} \qquad (1.16.31)$$

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

$$\left[g_{ij}\right] = \left[\frac{\partial x^k}{\partial \Theta^i}\right]^{\mathrm{T}}\left[\frac{\partial x^k}{\partial \Theta^j}\right], \qquad \left[g^{ij}\right] = \left[\frac{\partial \Theta^i}{\partial x^k}\right]\left[\frac{\partial \Theta^j}{\partial x^k}\right]^{\mathrm{T}} \qquad (1.16.32)$$

and, from 1.10.16a,b,

$$\det\left[g_{ij}\right] = \left(\det\left[\frac{\partial x^i}{\partial \Theta^j}\right]\right)^2 = J^2, \qquad \det\left[g^{ij}\right] = \left(\det\left[\frac{\partial \Theta^i}{\partial x^j}\right]\right)^2 = \frac{1}{J^2} \qquad (1.16.33)$$

These determinants play an important role, and are denoted by *g*:

$$g = \det\left[g_{ij}\right] = \frac{1}{\det\left[g^{ij}\right]}, \qquad \sqrt{g} = J \qquad (1.16.34)$$

<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 $|\mathbf{g}_i| = \sqrt{\mathbf{g}_i \cdot \mathbf{g}_i}$, $|\mathbf{g}^i| = \sqrt{\mathbf{g}^i \cdot \mathbf{g}^i}$, :

$$\hat{\mathbf{g}}_i = \frac{\mathbf{g}_i}{|\mathbf{g}_i|} = \frac{\mathbf{g}_i}{\sqrt{g_{ii}}}, \qquad \hat{\mathbf{g}}^i = \frac{\mathbf{g}^i}{|\mathbf{g}^i|} = \frac{\mathbf{g}^i}{\sqrt{g^{ii}}} \quad \text{(no sum)} \tag{1.16.35}$$

The lengths of the covariant base vectors are denoted by h and are called the **scale factors**:

$$h_i = |\mathbf{g}_i| = \sqrt{g_{ii}} \quad \text{(no sum)} \tag{1.16.36}$$

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

$$(\Delta s)^2 = d\mathbf{x} \cdot d\mathbf{x} = \left(d\Theta^i \mathbf{g}_i\right) \cdot \left(d\Theta^j \mathbf{g}_j\right) = g_{ij} d\Theta^i d\Theta^j \tag{1.16.38}$$

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.

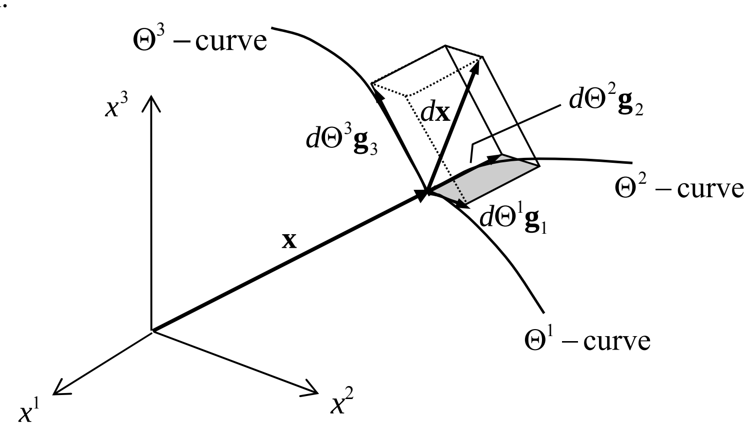


Figure 1.16.7: a line element in space

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.

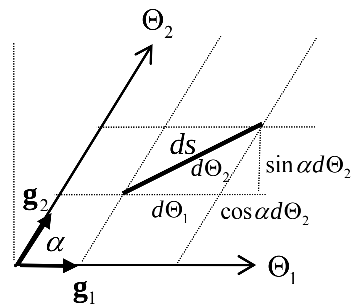


Figure 1.16.8: Length of a line element

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

$$\begin{aligned}(\Delta s)^2 &= h_1^2 d\Theta_1^2 + h_2^2 d\Theta_2^2 + h_3^2 d\Theta_3^2 \\[1em] \Delta S_1 &= h_2h_3\Delta\Theta_2\Delta\Theta_3 \\ \Delta S_2 &= h_3h_1\Delta\Theta_3\Delta\Theta_1 \\ \Delta S_3 &= h_1h_2\Delta\Theta_1\Delta\Theta_2 \\[1em] \Delta V &= h_1h_2h_3\Delta\Theta_1\Delta\Theta_2\Delta\Theta_3\end{aligned} \tag{1.16.46}$$

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:

$$x^1 = \Theta^1 \cos\Theta^2 \qquad \Theta^1 = \sqrt{\left(x^1\right)^2 + \left(x^2\right)^2}$$
$$x^2 = \Theta^1 \sin\Theta^2 , \qquad \Theta^2 = \tan^{-1}\left(x^2 / x^1\right)$$
$$x^3 = \Theta^3 \qquad \Theta^3 = x^3$$

with

$$\Theta^1 \geq 0, \quad 0 \leq \Theta^2 < 2\pi, \quad -\infty < \Theta^3 < \infty$$

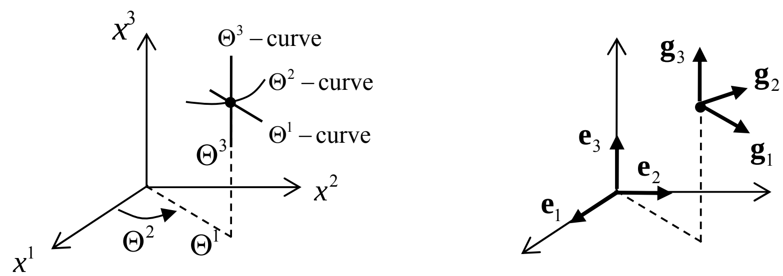


Figure 1.16.9: Cylindrical Coordinates

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}

$$\begin{aligned} h_1 &= \left|\mathbf{g}_1\right| = 1 \quad (=1) & \hat{\mathbf{g}}_1 &= \mathbf{g}_1 \quad (= \mathbf{e}_r) \\ h_2 &= \left|\mathbf{g}_2\right| = \Theta^1 (= r) & \hat{\mathbf{g}}_2 &= \frac{\mathbf{g}_2}{\Theta^1} (= \mathbf{e}_\theta) \\ h_3 &= \left|\mathbf{g}_3\right| = 1 \quad (=1) & \hat{\mathbf{g}}_3 &= \mathbf{g}_3 \quad (= \mathbf{e}_z) \end{aligned}$$

The line, surface and volume elements are

Metric: $$(\Delta s)^2 = \left(d\Theta^1\right)^2 + \left(\Theta^1 d\Theta^2\right)^2 + \left(d\Theta^3\right)^2 \quad \left(= dr^2 + (r d\theta)^2 + dz^2\right)$$

$$\Delta S_1 = \Theta^1 \Delta\Theta^2 \Delta\Theta^3$$

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

$$\Theta_1 \geq 0, \quad 0 \leq \Theta^2 \leq \pi, \quad 0 \leq \Theta^3 < 2\pi$$

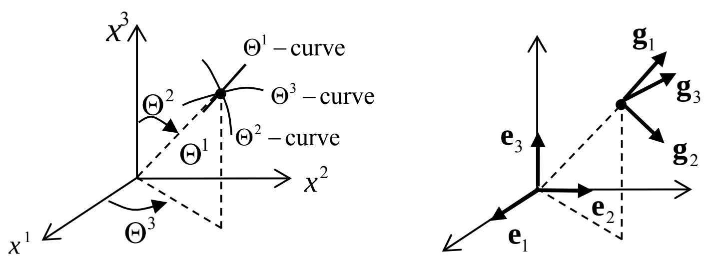


Figure 1.16.10: Spherical Coordinates

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:
$$(\Delta s)^2 = \left(d\Theta^1\right)^2 + \left(\Theta^1 d\Theta^2\right)^2 + \left(\Theta^1 \sin\Theta^2 d\Theta^3\right)^2$$
$$\left(= dr^2 + (rd\theta)^2 + (r\sin\theta d\phi)^2\right)$$

Surface Element:
$$
\begin{aligned}
\Delta S_1 &= \left(\Theta^1\right)^2 \sin\Theta^2 \Delta\Theta^2 \Delta\Theta^3 \\
\Delta S_2 &= \Theta^1 \sin\Theta^2 \Delta\Theta^3 \Delta\Theta^1 \\
\Delta S_3 &= \Theta^1 \Delta\Theta^1 \Delta\Theta^2
\end{aligned}
$$

Volume Element:
$$\Delta V = \left(\Theta^1\right)^2 \sin\Theta^2 \Delta\Theta^1 \Delta\Theta^2 \Delta\Theta^3 \quad \left(= r^2 \sin\theta \Delta r \Delta\theta \Delta\phi\right)$$

1.16.10 Vectors in Curvilinear Coordinates

A vector can now be represented in terms of *either* basis:

$$\mathbf{u} = u_i\left(\Theta^1, \Theta^2, \Theta^3\right)\mathbf{g}^i = u^i\left(\Theta^1, \Theta^2, \Theta^3\right)\mathbf{g}_i \tag{1.16.47}$$

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}:

$$\mathbf{u} \cdot \mathbf{g}_i = u_i, \quad \mathbf{u} \cdot \mathbf{g}^i = u^i \tag{1.16.48}$$

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}

$$u^i = g^{ij} u_j, \quad u_i = g_{ij} u^j \tag{1.16.49}$$

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,

$$\begin{aligned} \mathbf{u} &= u^i \mathbf{g}_i \\ &= \sum_{i=1}^{3} u^i h_i \hat{\mathbf{g}}_i \equiv u^{\langle i \rangle} \hat{\mathbf{g}}_i \end{aligned} \qquad (1.16.50)$$

and

$$\boxed{u^{\langle i \rangle} = h_i u^i = \sqrt{g_{ii}}\, u^i} \quad \text{(no sum)} \quad \textbf{Physical Components of a Vector} \qquad (1.16.51)$$

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}

$$\boxed{\mathbf{u} \cdot \mathbf{v} = u_i v^i = u^i v_i} \qquad \textbf{Dot Product of Two Vectors} \qquad (1.16.52)$$

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,

$$\begin{aligned} g &= [\mathbf{g}_1 \cdot \mathbf{g}_2 \times \mathbf{g}_3]^2 = \det[g_{ij}] \\ &= \frac{1}{[\mathbf{g}^1 \cdot \mathbf{g}^2 \times \mathbf{g}^3]^2} = \frac{1}{\det[g^{ij}]} \end{aligned} \qquad (1.16.53)$$

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],

$$\begin{array}{ll} (\mathbf{I})_{ij} = g_{ij} & \mathbf{I} = g_{ij} \left(\mathbf{g}^i \otimes \mathbf{g}^j\right) \\ (\mathbf{I})^{ij} = g^{ij} & \mathbf{I} = g^{ij} \left(\mathbf{g}_i \otimes \mathbf{g}_j\right) \\ (\mathbf{I})^{i}_{\cdot j} = \delta^i_j & \mathbf{I} = \delta^i_j \left(\mathbf{g}_i \otimes \mathbf{g}^j\right) = \mathbf{g}_i \otimes \mathbf{g}^i \\ (\mathbf{I})_{i}^{\cdot j} = \delta^j_i & \mathbf{I} = \delta^j_i \left(\mathbf{g}^i \otimes \mathbf{g}_j\right) = \mathbf{g}^i \otimes \mathbf{g}_i \end{array} \qquad (1.16.63)$$

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}	$A_i^{\cdot i} = A^i_{\cdot i}$
$\det \mathbf{A}$	$\varepsilon_{ijk} A_{i1} A_{j2} A_{k3}$	$\varepsilon_{ijk} A^i_{\cdot 1} A^j_{\cdot 2} A^k_{\cdot 3}$
$\mathbf{A}^{\mathrm{T}}$	$(\mathbf{A}^T)_{ij} = A_{ji}$	$(\mathbf{A}^{\mathrm{T}})_{ij} = A_{ji}$, $(\mathbf{A}^{\mathrm{T}})^{ij} = A^{ji}$ $(\mathbf{A}^{\mathrm{T}})^i_{\cdot j} = A_j^{\cdot i} \neq A^j_{\cdot i}$, $(\mathbf{A}^{\mathrm{T}})_i^{\cdot j} = A^j_{\cdot i} \neq A_j^{\cdot i}$

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.

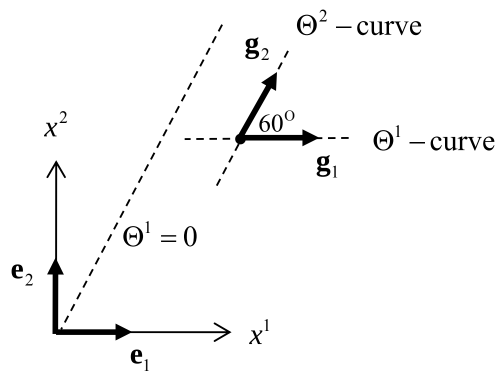


Figure 1.16.11: non-orthogoanl curvilinear coordinate system

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 $A^i_{\ j}$ 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 $\mathbf{A} = A^i_{\ j} \mathbf{g}^j \otimes \mathbf{g}_i$. 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 **u** can be obtained from $\mathbf{u} = u^i \mathbf{g}_i = \overline{u}^i \overline{\mathbf{g}}_i$ and $\mathbf{u} = u_i \mathbf{g}^i = \overline{u}_i \overline{\mathbf{g}}^i$:

$$\boxed{\begin{aligned} \overline{u}^i &= \frac{\partial \overline{\Theta}^i}{\partial \Theta^j}u^j, \qquad & u^i &= \frac{\partial \Theta^i}{\partial \overline{\Theta}^j}\overline{u}^j \\ \overline{u}_i &= \frac{\partial \Theta^j}{\partial \overline{\Theta}^i}u_j, \qquad & u_i &= \frac{\partial \overline{\Theta}^j}{\partial \Theta^i}\overline{u}_j \end{aligned}} \quad \textbf{Vector Transformation Rule} \tag{1.17.3}$$

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.

Note:

- 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,

$$\begin{aligned}
\overline{A}_{ij} &= \frac{\partial\Theta^m}{\partial\overline{\Theta}^i}\frac{\partial\Theta^n}{\partial\overline{\Theta}^j}A_{mn}, & A_{ij} &= \frac{\partial\overline{\Theta}^m}{\partial\Theta^i}\frac{\partial\overline{\Theta}^n}{\partial\Theta^j}\overline{A}_{mn} \\
\overline{A}^{ij} &= \frac{\partial\overline{\Theta}^i}{\partial\Theta^m}\frac{\partial\overline{\Theta}^j}{\partial\Theta^n}A^{mn}, & A^{ij} &= \frac{\partial\Theta^i}{\partial\overline{\Theta}^m}\frac{\partial\Theta^j}{\partial\overline{\Theta}^n}\overline{A}^{mn} \\
\overline{A}^i_{\cdot j} &= \frac{\partial\overline{\Theta}^i}{\partial\Theta^m}\frac{\partial\Theta^n}{\partial\overline{\Theta}^j}A^m_{\cdot n}, & A^i_{\cdot j} &= \frac{\partial\Theta^i}{\partial\overline{\Theta}^m}\frac{\partial\overline{\Theta}^n}{\partial\Theta^j}\overline{A}^m_{\cdot n} \\
\overline{A}_i^{\cdot j} &= \frac{\partial\overline{\Theta}^j}{\partial\Theta^n}\frac{\partial\Theta^m}{\partial\overline{\Theta}^i}A_m^{\cdot n}, & A_i^{\cdot j} &= \frac{\partial\Theta^j}{\partial\overline{\Theta}^n}\frac{\partial\overline{\Theta}^m}{\partial\Theta^i}\overline{A}_m^{\cdot n}
\end{aligned}$$

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

$$\frac{\partial\Theta^j}{\partial\overline{\Theta}^i} = \frac{\partial\overline{\Theta}^i}{\partial\Theta^j} \quad \to \quad Q_{ji} = Q_{ij}^{-1} \tag{1.17.6}$$

so the transformation is orthogonal, as expected. Also, as in Eqns. 1.5.11 and 1.5.13.

$$\begin{aligned}
u^i &= \frac{\partial\Theta^i}{\partial\overline{\Theta}^j}\overline{u}^j \quad \to \quad u_i = Q_{ij}u'_j \\
\overline{u}^i &= \frac{\partial\overline{\Theta}^i}{\partial\Theta^j}u^j \quad \to \quad u'_i = Q_{ij}^{-1}u_j = Q_{ji}u_j
\end{aligned} \tag{1.17.7}$$

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

$$\begin{aligned} \overline{(\Delta s)^2} &= \overline{g}_{ij} \overline{\Delta\Theta}^i \overline{\Delta\Theta}^j \\ &= \overline{\mathbf{g}_i \cdot \mathbf{g}_j} \overline{\Delta\Theta}^i \overline{\Delta\Theta}^j \\ &= \left(\frac{\partial \Theta^k}{\partial \overline{\Theta}^i} \mathbf{g}_k \cdot \frac{\partial \Theta^m}{\partial \overline{\Theta}^j} \mathbf{g}_m \right) \frac{\partial \overline{\Theta}^i}{\partial \Theta^p} \Delta\Theta^p \frac{\partial \overline{\Theta}^j}{\partial \Theta^q} \Delta\Theta^q \\ &= \delta_p^k \delta_q^m \left(\mathbf{g}_k \cdot \mathbf{g}_m\right) \Delta\Theta^p \Delta\Theta^q \\ &= g_{pq} \Delta\Theta^p \Delta\Theta^q \\ &= (\Delta s)^2 \end{aligned} \tag{1.17.15}$$

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}:

$$\frac{\partial g_{ij}}{\partial \Theta^k} = \Gamma_{ikj} + \Gamma_{jki} \qquad (1.18.9)$$

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)

$$\begin{aligned} \Gamma_{ijk} &= \frac{1}{2}\left(g_{jk,i} + g_{ki,j} - g_{ij,k}\right) \\ \Gamma^k_{ij} &= \frac{1}{2}g^{mk}\left(g_{jm,i} + g_{mi,j} - g_{ij,m}\right) \end{aligned} \qquad (1.18.10)$$

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)

$$\begin{aligned} g_{ij,k} &= \Gamma_{ikj} + \Gamma_{jki} \\ g^{ij}{}_{,k} &= -g^{im}\Gamma^j_{km} - g^{jm}\Gamma^i_{km} \end{aligned} \qquad (1.18.11)$$

Also, directly from 1.15.7, one has the relations

$$\frac{\partial g}{\partial g_{ij}} = g g^{ij}, \qquad \frac{\partial g}{\partial g^{ij}} = g g_{ij} \qquad (1.18.12)$$

and from these follow other useful relations, for example {▲Problem 4}

$$\Gamma_{ij}^{i} = \frac{\partial \log\left(\sqrt{g}\right)}{\partial \Theta^{j}} = \frac{1}{\sqrt{g}}\frac{\partial \sqrt{g}}{\partial \Theta^{j}} = J^{-1}\frac{\partial J}{\partial \Theta^{j}} \qquad (1.18.13)$$

and

$$\begin{aligned} \frac{\partial e_{ijk}}{\partial \Theta^{m}} &= \varepsilon_{ijk}\frac{\partial \sqrt{g}}{\partial \Theta^{m}} = \varepsilon_{ijk}\sqrt{g}\,\Gamma_{mn}^{n} = e_{ijk}\Gamma_{mn}^{n} \\ \frac{\partial e^{ijk}}{\partial \Theta^{m}} &= \varepsilon^{ijk}\frac{\partial\left(1/\sqrt{g}\right)}{\partial \Theta^{m}} = -\varepsilon^{ijk}\frac{1}{\sqrt{g}}\Gamma_{mn}^{n} = -e^{ijk}\Gamma_{mn}^{n} \end{aligned} \qquad (1.18.14)$$

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

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial \Theta^{j}} &= \frac{\partial v^{i}}{\partial \Theta^{j}}\mathbf{g}_{i} + v^{i}\frac{\partial \mathbf{g}_{i}}{\partial \Theta^{j}} & \qquad \frac{\partial \mathbf{v}}{\partial \Theta^{j}} &= \frac{\partial v_{i}}{\partial \Theta^{j}}\mathbf{g}^{i} + v_{i}\frac{\partial \mathbf{g}^{i}}{\partial \Theta^{j}} \\ &= \frac{\partial v^{i}}{\partial \Theta^{j}}\mathbf{g}_{i} + v^{i}\Gamma_{ij}^{k}\mathbf{g}_{k} \qquad \text{or} & &= \frac{\partial v_{i}}{\partial \Theta^{j}}\mathbf{g}^{i} - v_{i}\Gamma_{jk}^{i}\mathbf{g}^{k} \\ &\equiv v^{i}\big|_{j}\,\mathbf{g}_{i} & &\equiv v_{i}\big|_{j}\,\mathbf{g}^{i} \end{aligned} \qquad (1.18.15)$$

where

$$\boxed{\begin{aligned} v^{i}\,|_{j} &= v^{i}{}_{,j} + \Gamma_{kj}^{i}v^{k} \\ v_{i}\,|_{j} &= v_{i,j} - \Gamma_{ij}^{k}v_{k} \end{aligned}} \quad \textbf{Covariant Derivative of Vector Components} \qquad (1.18.16)$$

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

$$\begin{aligned} v_i\,|_{mn} &= B_{im}\,|_n \\ &= B_{im,n} - \Gamma^k_{in} B_{km} - \Gamma^k_{mn} B_{ik} \\ &= \left(v_{i,m} - \Gamma^k_{im} v_k\right)_{,n} - \Gamma^k_{in}\left(v_{k,m} - \Gamma^l_{km} v_l\right) - \Gamma^k_{mn}\left(v_{i,k} - \Gamma^l_{ik} v_l\right) \end{aligned} \tag{1.18.19}$$

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

$$v_i|_{mn} - v_i|_{nm} = R^{j}_{\cdot imn} v_j \qquad (1.18.20)$$

where **R** is the fourth-order **Riemann-Christoffel curvature tensor**, with (mixed) components

$$R^{j}_{\cdot imn} = \Gamma^j_{in,m} - \Gamma^j_{im,n} + \Gamma^k_{in}\Gamma^j_{km} - \Gamma^k_{im}\Gamma^j_{kn} \qquad (1.18.21)$$

Since the Christoffel symbols vanish in a Cartesian coordinate system, then so does $R^{j}_{\cdot imn}$. Again, any tensor that vanishes in one coordinate system must be zero in all coordinate systems, and so $R^{j}_{\cdot imn} = 0$, 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 $R^{j}_{\cdot imn} = 0$ 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

$$\nabla = \mathbf{g}^i \frac{\partial}{\partial \Theta^i} \qquad (1.18.22)$$

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

$$\boxed{\nabla\Phi \equiv \text{grad}\Phi = \frac{\partial \Phi}{\partial \Theta^i}\mathbf{g}^i} \qquad \textbf{Gradient of a Scalar} \qquad (1.18.23)$$

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 $\mathbf{g}^j$:

$$\boxed{\begin{aligned} \text{grad}\mathbf{u} = \frac{\partial \mathbf{u}}{\partial \Theta^j} \otimes \mathbf{g}^j &= u_i \,|_j \; \mathbf{g}^i \otimes \mathbf{g}^j \\ &= u^i \,|_j \; \mathbf{g}_i \otimes \mathbf{g}^j \end{aligned}} \qquad \textbf{Gradient of a Vector} \qquad (1.18.25)$$

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, $(\nabla \otimes \mathbf{u})^{\mathrm{T}} = \text{grad}\mathbf{u}$.

Again, one has for a space curve parameterised by s,

$$\frac{d\mathbf{u}}{ds} = \frac{\partial \mathbf{u}}{\partial \Theta^i} \tau^i = \frac{\partial \mathbf{u}}{\partial \Theta^i} \left(\boldsymbol{\tau} \cdot \mathbf{g}^i \right) = \left(\mathbf{g}^i \otimes \frac{\partial \mathbf{u}}{\partial \Theta^i} \right)^{\mathrm{T}} \cdot \boldsymbol{\tau} = \text{grad}\mathbf{u} \cdot \boldsymbol{\tau}$$

Similarly, from 1.18.18, the gradient of a second-order tensor is

$$\boxed{\begin{aligned} \text{grad}\mathbf{A} = \frac{\partial \mathbf{A}}{\partial \Theta^k} \otimes \mathbf{g}^k &= A^{ij} \,|_k \; \mathbf{g}_i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \\ &= A_{ij} \,|_k \; \mathbf{g}^i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \\ &= A_i^{\cdot j} \,|_k \; \mathbf{g}^i \otimes \mathbf{g}_j \otimes \mathbf{g}^k \\ &= A_{\cdot j}^{i} \,|_k \; \mathbf{g}_i \otimes \mathbf{g}^j \otimes \mathbf{g}^k \end{aligned}} \qquad \textbf{Gradient of a Tensor} \qquad (1.18.26)$$

The Divergence

From 1.14.9, the divergence of a vector is {▲Problem 6}

$$\boxed{\text{div}\mathbf{u} = \text{grad}\,\mathbf{u} : \mathbf{I} = u^i \,|_i \qquad \left(= \frac{\partial \mathbf{u}}{\partial \Theta^j} \cdot \mathbf{g}^j \right)} \qquad \textbf{Divergence of a Vector} \qquad (1.18.27)$$

This is equivalent to the divergence operation involving the Nabla operator, $\text{div}\mathbf{u} = \nabla \cdot \mathbf{u}$. An alternative expression can be obtained from 1.18.13 {▲Problem 7},

$$\operatorname{div}\mathbf{u} = u^i\,|_i = \frac{1}{\sqrt{g}}\frac{\partial\left(\sqrt{g}u^i\right)}{\partial\Theta^i} = J^{-1}\frac{\partial\left(Ju^i\right)}{\partial\Theta^i}$$

Similarly, using 1.14.12, the divergence of a second-order tensor is

$$\boxed{\begin{aligned}\operatorname{div}\mathbf{A} &= \operatorname{grad}\mathbf{A} : \mathbf{I} = A^{ij}\,|_j\;\mathbf{g}_i \qquad \left(= \frac{\partial \mathbf{A}}{\partial\Theta^j}\mathbf{g}^j\right)\\ &= A_i^{\cdot j}\,|_j\;\mathbf{g}^i\end{aligned}}$$

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, $\operatorname{div}\mathbf{A} = \nabla\cdot\mathbf{A}^{\mathrm{T}}$.

The Curl

The curl of a vector is defined by {▲Problem 8}

$$\boxed{\operatorname{curl}\mathbf{u} = \nabla\times\mathbf{u} = \mathbf{g}^k\times\frac{\partial\mathbf{u}}{\partial\Theta^k} = e^{ijk}u_j\,|_i\;\mathbf{g}_k = e^{ijk}\frac{\partial u_j}{\partial\Theta^i}\mathbf{g}_k}$$

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}

$$\begin{aligned}\operatorname{div}\mathbf{u} &= u^i\,|_i = u_i\,|^i\\ \operatorname{curl}\mathbf{u} &= e^{ijk}u_j\,|_i\;\mathbf{g}_k = e_{ijk}u^j\,|^i\;\mathbf{g}^k\end{aligned}.$$

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	$\text{grad}\phi,\ \nabla\phi = \partial\phi / \partial x_i$	$\phi_{,i} = \phi\|_i \equiv \partial\phi / \partial\Theta^i$
	of a vector field	$\text{grad}\mathbf{u} = \partial u_i / \partial x_j$	$u^i\|_j$
	of a tensor field	$\text{grad}\mathbf{T} = \partial T_{ij} / \partial x_k$	$T^{ij}\|_k$
Divergence	of a vector field	$\text{div}\mathbf{u},\ \nabla\cdot\mathbf{u} = \partial u_i / \partial x_i$	$u^i\|_i$
	of a tensor field	$\text{div}\mathbf{T} = \partial T_{ij} / \partial x_j$	$T^{ij}\|_j$
Curl	of a vector field	$\text{curl}\mathbf{u},\ \nabla\times\mathbf{u} = \varepsilon_{ijk}\partial u_j / \partial x_i$	$e^{ijk}u_j\|_i$

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}

$$\text{grad}(\alpha\mathbf{v}) = \alpha\text{grad}\mathbf{v} + \mathbf{v}\otimes\text{grad}\alpha$$
$$\text{div}(\mathbf{vA}) = \mathbf{v}\cdot\text{div}\mathbf{A} + \mathbf{A} : \text{grad}\mathbf{v}$$

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:

$$\frac{\partial\Phi}{\partial\mathbf{A}} = \frac{\partial\Phi}{\partial A_{ij}}\mathbf{g}_i\otimes\mathbf{g}_j = \frac{\partial\Phi}{\partial A_i^{\cdot j}}\mathbf{g}_i\otimes\mathbf{g}^j = \cdots$$
$$\frac{\partial\mathbf{B}}{\partial\mathbf{A}} = \frac{\partial B_{ij}}{\partial A_{mn}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_m\otimes\mathbf{g}_n = \cdots$$
$$\frac{\partial\mathbf{A}}{\partial\mathbf{A}} = \frac{\partial A_{ij}}{\partial A_{mn}}\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_m\otimes\mathbf{g}_n = \delta_i^m\delta_j^n\mathbf{g}^i\otimes\mathbf{g}^j\otimes\mathbf{g}_m\otimes\mathbf{g}_n$$
$$= \mathbf{g}^m\otimes\mathbf{g}^n\otimes\mathbf{g}_m\otimes\mathbf{g}_n$$

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}

$$\Gamma_{ii}^{k} = \Gamma_{ij}^{k}\Big|_{i=j} = -\frac{h_i}{h_k^2}\frac{\partial h_i}{\partial \Theta^k}, \quad i \neq k \qquad \text{(no sum)} \tag{1.18.35}$$

From the symmetry condition (see Eqn. 1.18.4), only 15 of the 21 non-zero symbols are distinct:

$$\begin{aligned}
&\Gamma_{11}^{1}, \Gamma_{12}^{1} = \Gamma_{21}^{1}, \Gamma_{13}^{1} = \Gamma_{31}^{1}, \Gamma_{22}^{1}, \Gamma_{33}^{1} \\
&\Gamma_{11}^{2}, \Gamma_{12}^{2} = \Gamma_{21}^{2}, \Gamma_{22}^{2}, \Gamma_{23}^{2} = \Gamma_{32}^{2}, \Gamma_{33}^{2} \\
&\Gamma_{11}^{3}, \Gamma_{13}^{3} = \Gamma_{31}^{3}, \Gamma_{22}^{3}, \Gamma_{23}^{3} = \Gamma_{32}^{3}, \Gamma_{33}^{3}
\end{aligned}$$

Note also that these are related to each other through the relation between (1.18.33, 1.18.35), i.e.

$$\Gamma_{ii}^{k} = -\frac{h_i^2}{h_k^2}\Gamma_{ik}^{i}, \quad i \neq k \qquad \text{(no sum)}$$

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

$$\operatorname{grad}\mathbf{v} = v^i\big|_j \mathbf{g}_i \otimes \mathbf{g}^j = \frac{1}{h_j^2} v^i\big|_j \mathbf{g}_i \otimes \mathbf{g}_j \qquad (\text{no sum over } h_j) \tag{1.18.37}$$

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}

$$\operatorname{div}\mathbf{v} = \frac{\partial v_i}{\partial \Theta^i} + v_k \Gamma^i_{ki} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial\left(v^{\langle 1\rangle} h_2 h_3\right)}{\partial \Theta^1} + \frac{\partial\left(v^{\langle 2\rangle} h_1 h_3\right)}{\partial \Theta^2} + \frac{\partial\left(v^{\langle 3\rangle} h_1 h_2\right)}{\partial \Theta^3}\right] \qquad (1.18.39)$$

The Curl of a Vector

From §1.16.10, the permutation symbol in orthogonal curvilinear coordinates reduceS to

$$e^{ijk} = \frac{1}{h_1 h_2 h_3}\varepsilon^{ijk} \qquad (1.18.40)$$

where $\varepsilon^{ijk} = \varepsilon_{ijk}$ is the Cartesian permutation symbol. From the definition 1.18.29, the curl of a vector is then

$$\begin{aligned}
\operatorname{curl}\mathbf{v} &= \frac{1}{h_1 h_2 h_3}\varepsilon_{ijk}\frac{\partial v_j}{\partial \Theta^i}\mathbf{g}_k = \frac{1}{h_1 h_2 h_3}\left\{\left[\frac{\partial v_2}{\partial \Theta^1} - \frac{\partial v_1}{\partial \Theta^2}\right]\mathbf{g}_3 + \cdots\right\} \\
&= \frac{1}{h_1 h_2 h_3}\left\{\left[\frac{\partial\left(v^2 h_2^2\right)}{\partial \Theta^1} - \frac{\partial\left(v^1 h_1^2\right)}{\partial \Theta^2}\right]\mathbf{g}_3 + \cdots\right\} \\
&= \frac{1}{h_1 h_2 h_3}\left\{\left[\frac{\partial\left(v^{\langle 2\rangle} h_2\right)}{\partial \Theta^1} - \frac{\partial\left(v^{\langle 1\rangle} h_1\right)}{\partial \Theta^2}\right]h_3\hat{\mathbf{g}}_3 + \cdots\right\}
\end{aligned} \qquad (1.18.41)$$

or

$$\operatorname{curl}\mathbf{v} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} h_1\hat{\mathbf{g}}_1 & h_2\hat{\mathbf{g}}_2 & h_3\hat{\mathbf{g}}_3 \\ \dfrac{\partial}{\partial \Theta^1} & \dfrac{\partial}{\partial \Theta^2} & \dfrac{\partial}{\partial \Theta^3} \\ h_1 v^{\langle 1\rangle} & h_2 v^{\langle 2\rangle} & h_3 v^{\langle 3\rangle} \end{vmatrix} \qquad (1.18.42)$$

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:

$$\operatorname{div}\mathbf{v} = \left[\frac{\partial v_{\langle 1\rangle}}{\partial\Theta^1} + \frac{v_{\langle 1\rangle}}{\Theta^1} + \frac{1}{\Theta^1}\frac{\partial v_{\langle 2\rangle}}{\partial\Theta^2} + \frac{\partial v_{\langle 3\rangle}}{\partial\Theta^3}\right]$$

$$\operatorname{curl}\mathbf{v} = \frac{1}{\Theta^1}\begin{vmatrix}\hat{\mathbf{g}}_1 & \Theta^1\hat{\mathbf{g}}_2 & \hat{\mathbf{g}}_3 \\ \dfrac{\partial}{\partial\Theta^1} & \dfrac{\partial}{\partial\Theta^2} & \dfrac{\partial}{\partial\Theta^3} \\ v^{\langle 1\rangle} & \Theta^1 v^{\langle 2\rangle} & v^{\langle 3\rangle}\end{vmatrix}$$

$$\begin{aligned}\operatorname{grad}\mathbf{v} &= \frac{\partial v_{\langle 1\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_1 + \frac{\partial v_{\langle 2\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_1 + \frac{\partial v_{\langle 3\rangle}}{\partial\Theta^1}\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_1 \\ &+ \frac{1}{\Theta^1}\left(\frac{\partial v_{\langle 1\rangle}}{\partial\Theta^2} - v_{\langle 2\rangle}\right)\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_2 + \frac{1}{\Theta^1}\left(\frac{\partial v_{\langle 2\rangle}}{\partial\Theta^2} + v_{\langle 1\rangle}\right)\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_2 \\ &+ \frac{1}{\Theta^1}\left(\frac{\partial v_{\langle 3\rangle}}{\partial\Theta^2}\right)\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_2 + \frac{\partial v_{\langle 1\rangle}}{\partial\Theta^3}\hat{\mathbf{g}}_1\otimes\hat{\mathbf{g}}_3 + \frac{\partial v_{\langle 2\rangle}}{\partial\Theta^3}\hat{\mathbf{g}}_2\otimes\hat{\mathbf{g}}_3 + \frac{\partial v_{\langle 3\rangle}}{\partial\Theta^3}\hat{\mathbf{g}}_3\otimes\hat{\mathbf{g}}_3\end{aligned}$$

The Divergence of a tensor {▲Problem 19} (see 1.14.19):

$$\begin{aligned}\operatorname{div}\mathbf{A} &= \left\{\frac{\partial A^{\langle 11\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 12\rangle}}{\partial\Theta^2} + \frac{\partial A^{\langle 13\rangle}}{\partial\Theta^3} + \frac{A^{\langle 11\rangle} - A^{\langle 22\rangle}}{\Theta^1}\right\}\hat{\mathbf{g}}_1 \\ &+ \left\{\frac{\partial A^{\langle 21\rangle}}{\partial\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 22\rangle}}{\partial\Theta^2} + \frac{\partial A^{\langle 23\rangle}}{\partial\Theta^3} + \frac{A^{\langle 21\rangle} + A^{\langle 12\rangle}}{\Theta^1}\right\}\hat{\mathbf{g}}_2 \\ &+ \left\{\frac{\partial A^{\langle 31\rangle}}{\partial\Theta^1} + \frac{A^{\langle 31\rangle}}{\Theta^1} + \frac{1}{\Theta^1}\frac{\partial A^{\langle 32\rangle}}{\partial\Theta^2} + \frac{\partial A^{\langle 33\rangle}}{\partial\Theta^3}\right\}\hat{\mathbf{g}}_3\end{aligned}$$

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. $\Gamma^k_{ij} = \Gamma^k_{ji}$, and that it is explicitly given by $\Gamma^k_{ij} = \dfrac{\partial\mathbf{g}_i}{\partial\Theta^j}\cdot\mathbf{g}^k$.

2 Consider the scalar-valued function $\Phi = (\mathbf{Au})\cdot\mathbf{v} = A_{ij}u^i v^j$. By taking the gradient of this function, and using the relation for the covariant derivative of **A**, i.e. $A_{ij}|_k = A_{ij,k} - \Gamma^m_{ik}A_{mj} - \Gamma^m_{jk}A_{im}$, 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 $\mathrm{div}\mathbf{u} = \mathrm{grad}\,\mathbf{u} : \mathbf{I} = u^i|_i$.

7 Derive the expression $\mathrm{div}\mathbf{u} = \left(1/\sqrt{g}\right)\partial\left(\sqrt{g}u^i\right)/\partial\Theta^i$

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

$$\mathrm{curl}(\mathbf{u}\times\mathbf{v}) = \begin{vmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 \\ \dfrac{\partial}{\partial\Theta^1} + \Gamma^k_{1k} & \dfrac{\partial}{\partial\Theta^2} + \Gamma^k_{2k} & \dfrac{\partial}{\partial\Theta^3} + \Gamma^k_{3k} \\ \left(u^2v^3 - u^3v^2\right) & -\left(u^1v^3 - u^3v^1\right) & \left(u^1v^2 - u^2v^1\right) \end{vmatrix}.$$

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) $\mathrm{grad}(\alpha\mathbf{v}) = \alpha\mathrm{grad}\mathbf{v} + \mathbf{v}\otimes\mathrm{grad}\alpha$, (ii) $\mathrm{div}(\mathbf{v}\mathbf{A}) = \mathbf{v}\cdot\mathrm{div}\mathbf{A} + \mathbf{A}:\mathrm{grad}\mathbf{v}$
[Hint: you might want to use the relation $\mathbf{a}\mathbf{T}\cdot\mathbf{b} = \mathbf{T}:(\mathbf{a}\otimes\mathbf{b})$ for the second of these.]

12 Derive the relation $\partial(\mathrm{tr}\mathbf{A})/\partial\mathbf{A} = \mathbf{I}$ 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 $\mathrm{grad}\Phi$.
(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 $\mathrm{div}\mathbf{u}$, $\mathrm{div}\mathbf{v}$
   (iii) Evaluate $\mathrm{curl}\mathbf{u}$, $\mathrm{curl}\mathbf{v}$
   (iv) Evaluate $\mathrm{grad}\mathbf{u}$, $\mathrm{grad}\mathbf{v}$
(g) Verify the vector identities

$$\begin{aligned}
\operatorname{div}(\Phi \mathbf{u}) &= \Phi \operatorname{div} \mathbf{u} + \operatorname{grad} \Phi \cdot \mathbf{u} \\
\operatorname{curl}(\Phi \mathbf{u}) &= \Phi \operatorname{curl} \mathbf{u} + \operatorname{grad} \Phi \times \mathbf{u} \\
\operatorname{div}(\mathbf{u} \times \mathbf{v}) &= \mathbf{v} \cdot \operatorname{curl} \mathbf{u} - \mathbf{u} \cdot \operatorname{curl} \mathbf{v} \\
\operatorname{curl}(\operatorname{grad} \Phi) &= \mathbf{o} \\
\operatorname{div}(\operatorname{curl} \mathbf{u}) &= 0
\end{aligned}$$

(h) Verify the identities

$$\begin{aligned}
\operatorname{grad}(\Phi \mathbf{v}) &= \Phi \operatorname{grad} \mathbf{v} + \mathbf{v} \otimes \operatorname{grad} \Phi \\
\operatorname{grad}(\mathbf{u} \cdot \mathbf{v}) &= (\operatorname{grad} \mathbf{u})^{\mathrm{T}} \mathbf{v} + (\operatorname{grad} \mathbf{v})^{\mathrm{T}} \mathbf{u} \\
\operatorname{div}(\mathbf{u} \otimes \mathbf{v}) &= (\operatorname{grad} \mathbf{u}) \mathbf{v} + (\operatorname{div} \mathbf{v}) \mathbf{u} \\
\operatorname{curl}(\mathbf{u} \times \mathbf{v}) &= \mathbf{u} \operatorname{div} \mathbf{v} - \mathbf{v} \operatorname{div} \mathbf{u} + (\operatorname{grad} \mathbf{u}) \mathbf{v} - (\operatorname{grad} \mathbf{v}) \mathbf{u}
\end{aligned}$$

(i) Consider the tensor field

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & -\Theta^2 \end{bmatrix} (\mathbf{g}^i \otimes \mathbf{g}^j)$$

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

$$\frac{\partial}{\partial \Theta^i} (\mathbf{g}_i \cdot \mathbf{g}_j) = 0, \quad i \neq j$$

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.

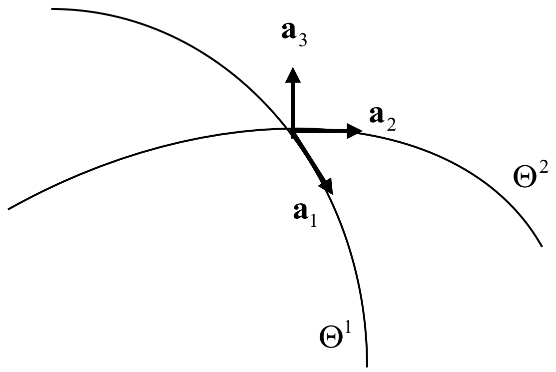


Figure 1.19.1: Geometry of the Curved Surface

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

$$e_{\alpha\beta} \equiv \mathbf{a}_\alpha \cdot \mathbf{a}_\beta \times \mathbf{a}_3 = \varepsilon_{\alpha\beta}\sqrt{g}, \qquad e^{\alpha\beta} \equiv \mathbf{a}^\alpha \cdot \mathbf{a}^\beta \times \mathbf{a}^3 = \varepsilon^{\alpha\beta}\frac{1}{\sqrt{g}} \qquad (1.19.3)$$

where $\varepsilon_{\alpha\beta} = \varepsilon^{\alpha\beta}$ is the Cartesian permutation symbol, $\varepsilon_{12} = +1$, $\varepsilon_{21} = -1$, and zero otherwise, with

$$e^{\alpha\beta} e_{\mu\eta} = \varepsilon^{\alpha\beta}\varepsilon_{\mu\eta}, \qquad e^{\alpha\beta} e_{\mu\eta} = \delta^\alpha_\mu \delta^\beta_\eta - \delta^\beta_\mu \delta^\alpha_\eta = e^{\beta\alpha} e_{\eta\mu} \qquad (1.19.4)$$

From 1.19.3,

$$\begin{aligned} \mathbf{a}_\alpha \times \mathbf{a}_\beta &= e_{\alpha\beta}\mathbf{a}^3 \\ \mathbf{a}^\alpha \times \mathbf{a}^\beta &= e^{\alpha\beta}\mathbf{a}_3 \end{aligned} \qquad (1.19.5)$$

and so

$$\mathbf{a}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\sqrt{g}} \qquad (1.19.6)$$

The cross product of surface vectors, that is, vectors with component in the normal ($\mathbf{g}_3$) direction zero, can be written as

$$\begin{aligned} \mathbf{u} \times \mathbf{v} = e_{\alpha\beta} u^\alpha v^\beta \mathbf{a}^3 &= \sqrt{g}\begin{vmatrix} u^1 & u^2 \\ v^1 & v^2 \end{vmatrix} \mathbf{a}^3 \\ = e^{\alpha\beta} u_\alpha v_\beta \mathbf{a}_3 &= \frac{1}{\sqrt{g}}\begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \mathbf{a}_3 \end{aligned} \qquad (1.19.7)$$

The Metric and Surface elements

Considering a line element lying within the surface, so that $\Theta^3 = 0$, the metric for the surface is

$$(\Delta s)^2 = d\mathbf{s} \cdot d\mathbf{s} = \left(d\Theta^\alpha \mathbf{a}_\alpha\right) \cdot \left(d\Theta^\beta \mathbf{a}_\beta\right) = g_{\alpha\beta} d\Theta^\alpha d\Theta^\beta \qquad (1.19.8)$$

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

$$\Delta S = \sqrt{g}\,\Delta\Theta^1 \Delta\Theta^2 \qquad (1.19.9)$$

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

$$d\mathbf{a}_3 = \mathbf{a}_{3,1} d\Theta^1 = -K_1^\gamma d\Theta^1 \mathbf{a}_\gamma \tag{1.19.24}$$

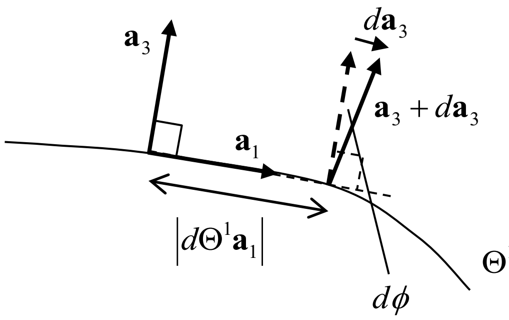


Figure 1.19.2: Curvature of the Surface

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 $|d\mathbf{a}_3|$ 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

$$\frac{d\varphi}{ds} = \frac{\left|-K_1^2 d\Theta^1 \mathbf{a}_2\right|}{\left|d\Theta^1 \mathbf{a}_1\right|} = \left|K_1^2\right| \frac{\left|\mathbf{a}_2\right|}{\left|\mathbf{a}_1\right|} \tag{1.19.26}$$

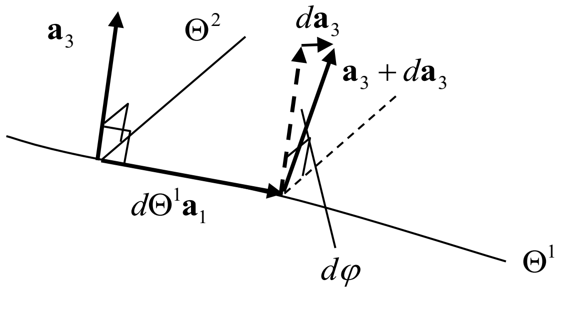


Figure 1.19.3: Twisting over the Surface

When $|\mathbf{a}_1| = |\mathbf{a}_2|$, $\left|K_1^2\right|$ 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:

$$\begin{aligned} \mathrm{I}_{\mathbf{K}} &= K^{i}_{\cdot i} = K^{1}_{1} + K^{2}_{2} \\ \mathrm{III}_{\mathbf{K}} &= \det K^{i}_{\cdot j} = \begin{vmatrix} K^{1}_{1} & K^{1}_{2} \\ K^{2}_{1} & K^{2}_{2} \end{vmatrix} = K^{1}_{1}K^{2}_{2} - K^{1}_{2}K^{2}_{1} = \varepsilon_{\alpha\beta}K^{\alpha}_{1}K^{\beta}_{2} \end{aligned} \tag{1.19.27}$$

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

$$x^1 = a\sin\Theta^1\cos\Theta^2, \quad x^2 = a\sin\Theta^1\sin\Theta^2, \quad x^3 = a\cos\Theta^1$$

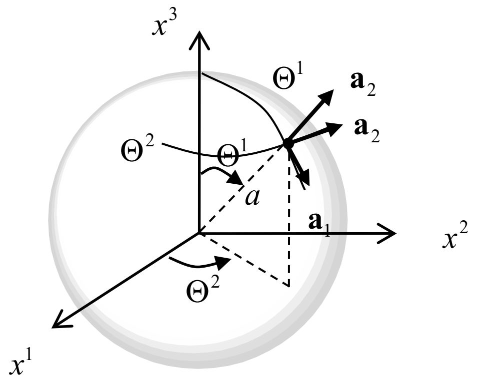


Figure 1.19.4: a spherical surface

Then, from the definitions 1.16.19, 1.16.27-28, 1.16.34, {▲Problem 2}

$$\begin{aligned} \mathbf{a}_1 &= +a\cos\Theta^1\cos\Theta^2\mathbf{e}_1 + a\cos\Theta^1\sin\Theta^2\mathbf{e}_2 - a\sin\Theta^1\mathbf{e}_3 \\ \mathbf{a}_2 &= -a\sin\Theta^1\sin\Theta^2\mathbf{e}_1 + a\sin\Theta^1\cos\Theta^2\mathbf{e}_2 \\ \mathbf{a}^1 &= \frac{1}{a^2}\mathbf{a}_1 \\ \mathbf{a}^2 &= \frac{1}{a^2\sin^2\Theta^1}\mathbf{a}_2 \\ g_{\alpha\beta} &= \begin{vmatrix} a^2 & 0 \\ 0 & a^2\sin^2\Theta^1 \end{vmatrix}, \qquad g = a^4\sin^2\Theta^1 \end{aligned} \tag{1.19.28}$$

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,

$$\left[K^\alpha_\beta\right] = \begin{bmatrix} -\frac{1}{a} & 0 \\ 0 & -\frac{1}{a} \end{bmatrix}, \qquad \left[K_{\alpha\beta}\right] = \begin{bmatrix} -a & 0 \\ 0 & -a\sin^2\Theta^1 \end{bmatrix} \tag{1.19.33}$$

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

$$
\begin{aligned}
v^\alpha|_\beta &= v^\alpha{}_{,\beta} + \Gamma^\alpha_{\gamma\beta} v^\gamma + \Gamma^\alpha_{3\beta} v^3 \\
v^\alpha|_3 &= v^\alpha{}_{,3} + \Gamma^\alpha_{\gamma 3} v^\gamma + \Gamma^\alpha_{33} v^3 \\
&= v^\alpha{}_{,3} + \Gamma^\alpha_{\gamma 3} v^\gamma \\
v^3|_\alpha &= v^3{}_{,\alpha} + \Gamma^3_{\gamma\alpha} v^\gamma + \Gamma^3_{3\alpha} v^3 \\
&= v^3{}_{,\alpha} + \Gamma^3_{\gamma\alpha} v^\gamma
\end{aligned}
\quad , \quad
\begin{aligned}
v_\alpha|_\beta &= v_{\alpha,\beta} - \Gamma^\gamma_{\alpha\beta} v_\gamma - \Gamma^3_{\alpha\beta} v_3 \\
v_\alpha|_3 &= v_{\alpha,3} - \Gamma^\gamma_{\alpha 3} v_\gamma - \Gamma^3_{\alpha 3} v_3 \\
&= v_{\alpha,3} - \Gamma^\gamma_{\alpha 3} v_\gamma \\
v_3|_\alpha &= v_{3,\alpha} - \Gamma^\gamma_{3\alpha} v_\gamma - \Gamma^3_{3\alpha} v_3 \\
&= v_{3,\alpha} - \Gamma^\gamma_{3\alpha} v_\gamma
\end{aligned}
\qquad (1.19.35)
$$

Define now a two-dimensional analogue of the three-dimensional covariant derivative through

$$
\begin{aligned}
v^\alpha\|_\beta &= v^\alpha{}_{,\beta} + \Gamma^\alpha_{\gamma\beta} v^\gamma \\
v_\alpha\|_\beta &= v_{\alpha,\beta} - \Gamma^\gamma_{\alpha\beta} v_\gamma
\end{aligned}
\qquad (1.19.36)
$$

so that, using 1.19.19, 1.19.21, the covariant derivative can be expressed as

$$
\begin{aligned}
v^\alpha|_\beta &= v^\alpha\|_\beta - K^\alpha_\beta v^3 \\
v_\alpha|_\beta &= v_\alpha\|_\beta - K_{\alpha\beta} v_3
\end{aligned}
\qquad (1.19.37)
$$

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

$$
\begin{aligned}
\mathbf{v}_{,\beta} &= v^i|_\beta \, \mathbf{a}_i \\
&= \left(v^\alpha\|_\beta - K^\alpha_\beta v^3\right)\mathbf{a}_\alpha + v^3|_\beta \, \mathbf{a}_3 \\
\mathbf{v}_{,\beta} &= v_i|_\beta \, \mathbf{a}^i \\
&= \left(v_\alpha\|_\beta - K_{\alpha\beta} v_3\right)\mathbf{a}^\alpha + v_3|_\beta \, \mathbf{a}^3
\end{aligned}
\qquad (1.19.38)
$$

From 1.18.25, the gradient of a surface vector is (using 1.19.21)

$$
\operatorname{grad} \mathbf{v} = \left(v_\alpha\|_\beta - K_{\alpha\beta} v_3\right)\mathbf{a}^\alpha \otimes \mathbf{a}^\beta + K^\gamma_\alpha v_\gamma \mathbf{a}^\alpha \otimes \mathbf{a}^3 \qquad (1.19.39)
$$

Tensors

The covariant derivatives of second order tensor components are given by 1.18.18. For example,

$$
\begin{aligned}
A^{ij}|_\gamma &= A^{ij}{}_{,\gamma} + \Gamma^i_{m\gamma} A^{mj} + \Gamma^j_{m\gamma} A^{im} \\
&= A^{ij}{}_{,\gamma} + \Gamma^i_{\lambda\gamma} A^{\lambda j} + \Gamma^i_{3\gamma} A^{3j} + \Gamma^j_{\lambda\gamma} A^{i\lambda} + \Gamma^j_{3\gamma} A^{i3}
\end{aligned}
\qquad (1.19.40)
$$

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,

$$\boldsymbol{\tau}_2 = \mathbf{a}_3 \times \boldsymbol{\tau} = \frac{d\Theta^\alpha}{ds}\mathbf{a}_3 \times \mathbf{a}_\alpha = e_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\mathbf{a}^\beta \qquad (1.19.60)$$

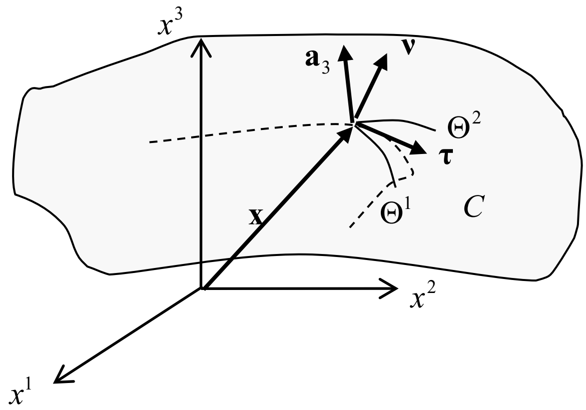


Figure 1.19.5: a curve lying on a surface

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}

$$\frac{d\boldsymbol{\tau}}{ds} = \left(\frac{d^2\Theta^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\right)\mathbf{a}_\gamma + K_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\mathbf{a}_3 \qquad (1.19.61)$$

Then

$$\frac{d\boldsymbol{\tau}}{ds} = \kappa_g \boldsymbol{\tau}_2 + \kappa_n \mathbf{a}_3 \qquad (1.19.62)$$

where

$$\begin{aligned} \kappa_g &= e_{\lambda\gamma}\frac{d\Theta^\lambda}{ds}\left(\frac{d^2\Theta^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds}\right) \\ \kappa_n &= K_{\alpha\beta}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds} \end{aligned} \qquad (1.19.63)$$

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

$$\kappa_n = \boldsymbol{\tau}\,\mathbf{K}\,\boldsymbol{\tau} \qquad (1.19.64)$$

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

$$\left[K_i^{\,j}\right] = \begin{bmatrix} K_1^1 & 0 \\ 0 & K_2^2 \end{bmatrix} = \begin{bmatrix} (\kappa_n)_{\max} & 0 \\ 0 & (\kappa_n)_{\min} \end{bmatrix} \qquad (1.19.65)$$

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

$$\frac{d^2\Theta^\gamma}{ds^2} + \Gamma_{\alpha\beta}^{\gamma}\frac{d\Theta^\alpha}{ds}\frac{d\Theta^\beta}{ds} = 0 \qquad (1.19.64)$$

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, $\Gamma_{\alpha 3}^{3} = \Gamma_{33}^{\alpha} = \Gamma_{33}^{3} = 0$.
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 $aa^{-1} = a^{-1}a = e$

<u>Examples</u>:

(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 $a, b \in G$.

<u>Examples</u>:

(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', $k = \{a \in G \mid f(a) = e'\}$

<u>Examples</u>

(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

<u>Examples</u>:

(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.

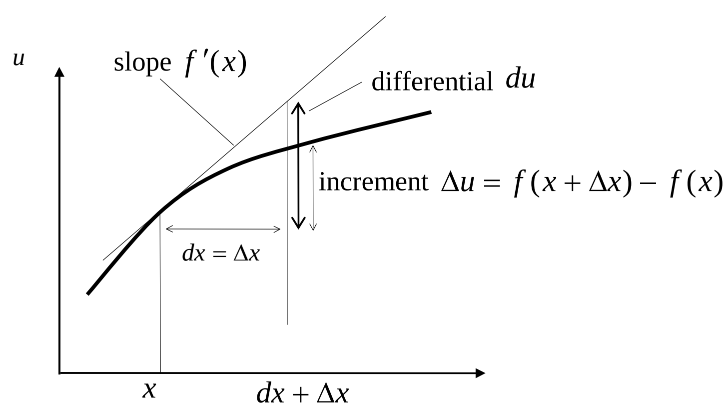


Figure 1.B.1: the differential

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, $(\Delta s)^2 \approx (\Delta x)^2 + (\Delta y)^2$.

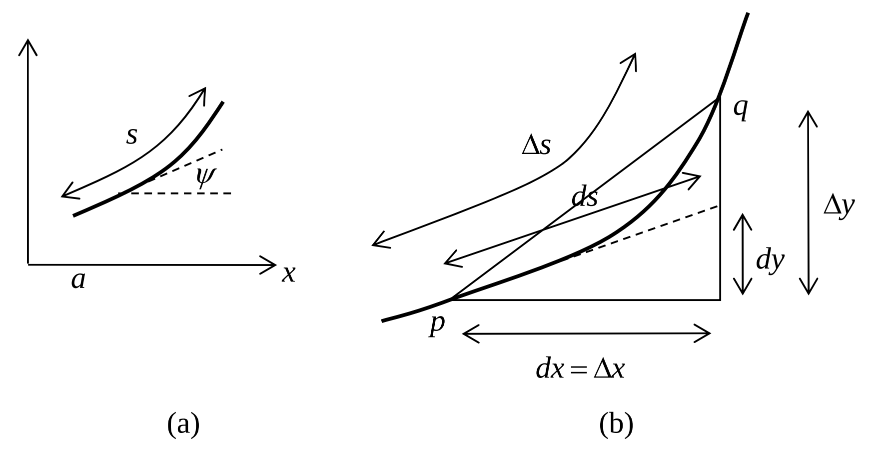


Figure 1.B.2: arc length

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,

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} \tag{1.B.8}$$

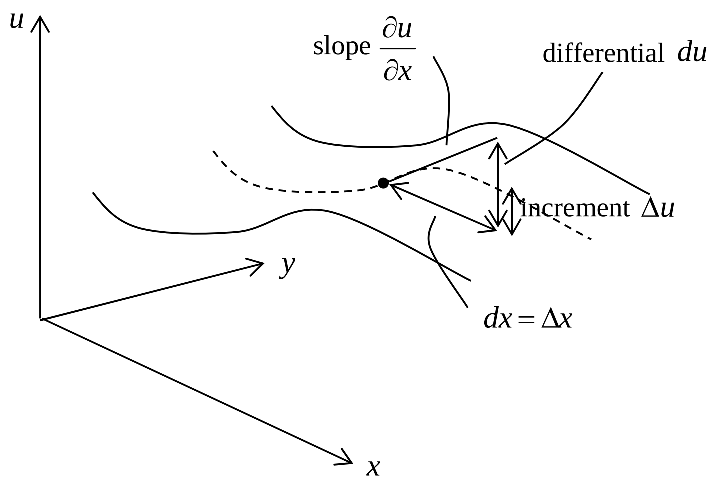


Figure 1.B.3: the partial derivative

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 $u = f[x(t), y(t)]$, 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 $x = x(t)$, $u = f[x(t), t]$, 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 $u = f[x(t)]$, 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**.

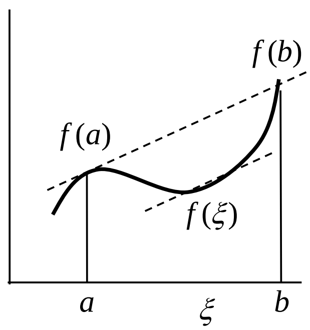


Figure 1.B.4: the mean value theorem

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**.

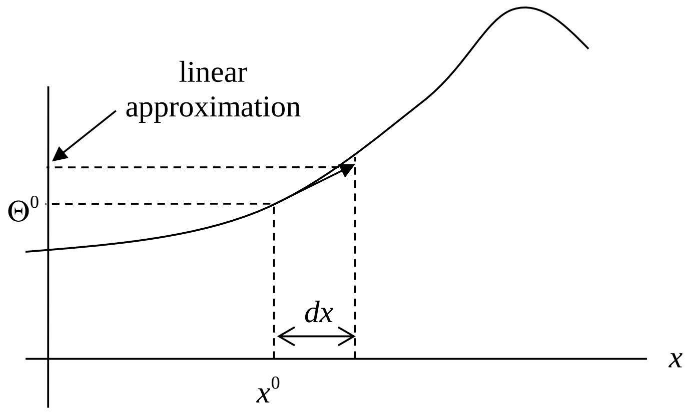


Figure 1.B.5: linear approximation