Ruhr-University Bochum
Faculty of Civil and Environmental Engineering
Institute of Mechanics

A Small Compendium on Vector and Tensor Algebra and Calculus

Klaus Hackl

Mehdi Goodarzi

2010

Contents

Foreword

A quick review of vector and tensor algebra, geometry and analysis is presented. The reader is supposed to have sufficient familiarity with the subject and the material is included as an entry point as well as a reference for the subsequence.

Chapter 1

Vectors and tensors algebra

Algebra is concerned with operations defined in sets with certain properties. Tensor and vector algebra deals with properties and operations in the set of tensors and vectors. Throughout this section together with algebraic aspects, we also consider geometry of tensors to obtain further insight.

1.1 Scalars and vectors

There are physical quantities that can be specified by a single real number, like temperature and speed. These are called *scalars*. Other quantities whose specification requires magnitude (a positive real number) and direction, for instance velocity and force, are called *vectors*. A typical illustration of a vector is an arrow which is a directed line segment. Given an appropriate unit, length of the arrow reflects magnitude of the vector. In the present context the precise terminology would be *Euclidean* vector or *geometric* vector, to acknowledge that in mathematics, vector has a broader meaning with which we are not concerned here. It is instructive to have a closer look at the meaning of direction of a vector.

Given a vector $\boldsymbol{v}$ and a directed line $\mathscr{L}$ (Fig. 1.1), projection of the vector onto the direction is a scalar $v_{\mathscr{L}} = \|\boldsymbol{v}\| \cos\theta$ in which θ is the angle between $\boldsymbol{v}$ and $\mathscr{L}$. Therefore, a vector can be thought as a function that assigns a scalar to each direction $\boldsymbol{v}(\mathscr{L}) = v_{\mathscr{L}}$. This notion can be naturally extended to understand tensors. A tensor (of second order) is a function that assigns vectors to directions $\boldsymbol{T}(\mathscr{L}) = \boldsymbol{T}_{\mathscr{L}}$ in the sense of projection. In other words the projection of tensor $\boldsymbol{T}$ on direction $\mathscr{L}$ is a vector like $\boldsymbol{T}_{\mathscr{L}}$. This looks rather abstract but its meaning is going to be clear in the sequel when we explain the *Cauchy's formula* in which the dot product of stress (tensor) and area (vector) yields traction force (vector).

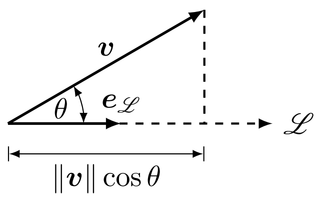

Fig. 1.1: Projection.

Notation 1. Vectors and tensors are denoted by bold-face letters like $\boldsymbol{A}$ and $\boldsymbol{v}$, and magnitude of a vector by $\|\boldsymbol{v}\|$ or simply by its normal-face v. In handwriting an underbar is used to denote tensors and vectors like $\underline{v}$.

1.2 Geometrical operations on vectors

Definition 2 (Equality of vectors)**.** Two vectors are equal if they have the same magnitude and direction.

Definition 3 (Summation of vectors)**.** The sum of every two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is a vector $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$ formed by placing initial point of $\boldsymbol{w}$ on terminal point of $\boldsymbol{v}$ and joining initial point of $\boldsymbol{v}$ and terminal point of $\boldsymbol{w}$ (Fig. 1.2(a,b)). This is equivalent to parallelogram law for vector addition as shown in Fig. 1.2(c). Subtraction of vectors $\boldsymbol{v} - \boldsymbol{w}$ is defined as $\boldsymbol{v} + (-\boldsymbol{w})$.

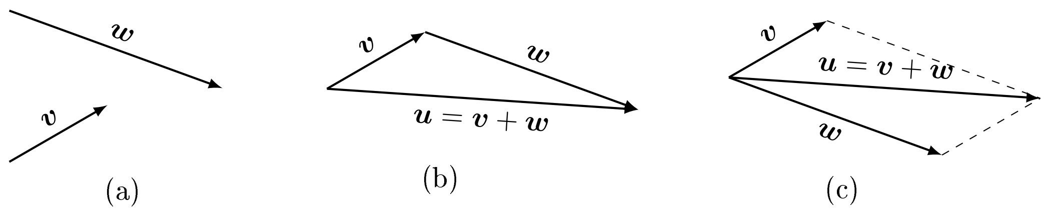

Fig. 1.2: Vector summation.

Definition 4 (Multiplication of a vector by a scalar)**.** For any real number (scalar) α and vector $\boldsymbol{v}$, their multiplication is a vector $\alpha\boldsymbol{v}$ whose magnitude is $|\alpha|$ times the magnitude of $\boldsymbol{v}$ and whose direction is the same as $\boldsymbol{v}$ if $\alpha > 0$ and opposite to $\boldsymbol{v}$ if $\alpha < 0$. If $\alpha = 0$ then $\alpha\boldsymbol{v} = \boldsymbol{0}$.

1.3 Fundamental properties

The set of geometric vectors V together with the just defined summation and multiplication operations is subject to the following properties

Property 5 (Vector summation)**.**

1.	$\boldsymbol{v} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{v} = \boldsymbol{v}$	Existence of additive identity
2.	$\boldsymbol{v} + (-\boldsymbol{v}) = (-\boldsymbol{v}) + \boldsymbol{v} = \boldsymbol{0}$	Existence of additive inverse
3.	$\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$	Commutative law
4.	$\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$	Associative law

Property 6 (Multiplication with scalars)**.**

1.	$\alpha\,(\boldsymbol{v} + \boldsymbol{w}) = \alpha\boldsymbol{v} + \alpha\boldsymbol{w}$	Distributive law
2.	$(\alpha + \beta)\,\boldsymbol{v} = \alpha\boldsymbol{v} + \beta\boldsymbol{v}$	Distributive law
3.	$\alpha\,(\beta\boldsymbol{v}) = (\alpha\beta)\,\boldsymbol{v}$	Associative law
4.	$1\,\boldsymbol{v} = \boldsymbol{v}$	Multiplication with scalar identity.

A vector can be rigorously defined as an element of a vector space.

Definition 7 (Real vector space)**.** The mathematical system $(V, +, \mathbb{R}, *)$ is called a real vector space or simply a vector space, where V is the set of geometrical vectors, $+$ is vector summation with the four properties mentioned, $\mathbb{R}$ is the set of real numbers furnished with summation and multiplication of numbers, and $*$ stands for multiplication of a vector with a real number (scalar) having the above mentioned properties.

The reader may wonder why this recent definition is needed. Remember that a quantity whose expression requires magnitude and direction is not necessarily a vector(!). A well known example is finite rotation which has both magnitude and direction but is not a vector because it does not follow the properties of vector summation, as defined before. Therefore, having magnitude and direction is not sufficient for a quantity to be identified as a vector, but it must also obey the rules of summation and multiplication with scalars.

Definition 8 (Unit vector). A unit vector is a vector of unit magnitude. For any vector $\boldsymbol{v}$ its unit vector is referred to by $\boldsymbol{e}_v$ or $\hat{\boldsymbol{v}}$ which is equal to $\hat{\boldsymbol{v}} = \boldsymbol{v}/\|\boldsymbol{v}\|$. One would say that the unit vector carries the information about direction. Therefore magnitude and direction as constituents of a vector are multiplicatively decomposed as $\boldsymbol{v} = v\hat{\boldsymbol{v}}$.

1.4 Vector decomposition

Writing a vector $\boldsymbol{v}$ as the sum of its components is called decomposition. Components of a vector are its projections onto the coordinate axes (Fig. 1.3)

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3\,, \tag{1.1}$$

where $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ are *unit basis vectors* of coordinate system, $\{v_1, v_2, v_3\}$ are *components* of $\boldsymbol{v}$ and $\{v_1\boldsymbol{e}_1, v_2\boldsymbol{e}_2, v_3\boldsymbol{e}_3\}$ are *component vectors* of $\boldsymbol{v}$.

Since a vector is specified by its components relative to a given coordinate system, it can be alternatively denoted by *array notation*

$$\boldsymbol{v} \equiv \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = (v_1\ v_2\ v_3)^T\,. \tag{1.2}$$

Using Pythagorean theorem the norm (magnitude) of a vector $\boldsymbol{v}$ in three-dimensional space can be written as

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}\,. \tag{1.3}$$

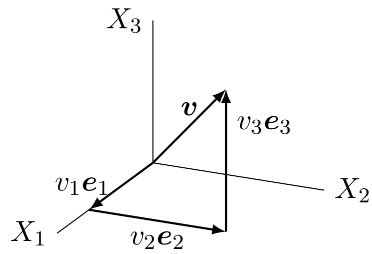

Fig. 1.3: Components.

Remark 9. Representation of a geometrical vector by its equivalent array (or tuple) is the underlying idea of analytic geometry which basically paves the way for an algebraic treatment of geometry.

1.5 Dot product of vectors

Dot product (also called inner or scalar product) of two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is defined as

$$\boldsymbol{v} \cdot \boldsymbol{w} = v\,w\,\cos\theta\,, \tag{1.4}$$

where θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$.

Property 10 (Dot product).

1.	$\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{v}$	Commutativity
2.	$\boldsymbol{u} \cdot (\alpha \boldsymbol{v} + \beta \boldsymbol{w}) = \alpha \boldsymbol{u} \cdot \boldsymbol{v} + \beta \boldsymbol{u} \cdot \boldsymbol{w}$	Linearity
3.	$\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$ and $\boldsymbol{v} \cdot \boldsymbol{v} = 0 \Leftrightarrow \boldsymbol{v} = \boldsymbol{0}$	Positivity.

Geometrical interpretation of dot product

Definition of dot product (1.4) can be restated as

$$\boldsymbol{v} \cdot \boldsymbol{w} = (v\hat{\boldsymbol{v}}) \cdot (w\hat{\boldsymbol{w}}) = v\,(\hat{\boldsymbol{v}} \cdot \boldsymbol{w}) = w\,(\boldsymbol{v} \cdot \hat{\boldsymbol{w}}) = vw\cos\theta\,, \tag{1.5}$$

which says the dot product of vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ equals projection of $\boldsymbol{w}$ on $\boldsymbol{v}$ direction times magnitude of $\boldsymbol{v}$ (see Fig. 1.1), or the other way around. Also, dividing leftmost and rightmost terms of (1.5) by vw gives

$$\hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{w}} = \cos\theta\,, \tag{1.6}$$

which says the dot product of unit vectors (standing for directions) determines the angle between them. A simple and yet interesting result is obtained for components of a vector $\boldsymbol{v}$ as

$$v_1 = \boldsymbol{v} \cdot \boldsymbol{e}_1\,, \quad v_2 = \boldsymbol{v} \cdot \boldsymbol{e}_2\,, \quad v_3 = \boldsymbol{v} \cdot \boldsymbol{e}_3. \tag{1.7}$$

Important results

1. In any *orthonormal* coordinate system including Cartesian coordinates it holds that

$$\begin{aligned} \boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = 1\,,\ \ \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = 0\,,\ \ \boldsymbol{e}_1 \cdot \boldsymbol{e}_3 = 0 \\ \boldsymbol{e}_2 \cdot \boldsymbol{e}_1 = 0\,,\ \ \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 = 1\,,\ \ \boldsymbol{e}_2 \cdot \boldsymbol{e}_3 = 0 \\ \boldsymbol{e}_3 \cdot \boldsymbol{e}_1 = 0\,,\ \ \boldsymbol{e}_3 \cdot \boldsymbol{e}_2 = 0\,,\ \ \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 = 1\,, \end{aligned} \tag{1.8}$$

2. therefore *

$$\boldsymbol{v} \cdot \boldsymbol{w} = v_1 w_1 + v_2 w_2 + v_3 w_3\,. \tag{1.9}$$

3. For any two nonzero vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ iff $\boldsymbol{a}$ is normal to $\boldsymbol{b}$.

4. Norm of a vector $\boldsymbol{v}$ can be obtained by

$$\|\boldsymbol{v}\| = (\boldsymbol{v} \cdot \boldsymbol{v})^{1/2}\,. \tag{1.10}$$

* Note that $\boldsymbol{v} \cdot \boldsymbol{w} = (v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3) \cdot (w_1\boldsymbol{e}_1 + w_2\boldsymbol{e}_2 + w_3\boldsymbol{e}_3) = v_1 w_1 \boldsymbol{e}_1 \cdot \boldsymbol{e}_1 + v_1 w_2 \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 + \cdots$.

1.6 Index notation

A typical tensorial calculation demands component-wise arithmetic derivations which is usually a tedious task. As an example let us try to expand the left-hand-side of equation (1.9)

$$\begin{aligned}\boldsymbol{v}\cdot\boldsymbol{w} &= (v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3)\cdot(w_1\boldsymbol{e}_1 + w_2\boldsymbol{e}_2 + w_3\boldsymbol{e}_3)\\ &= v_1w_1\boldsymbol{e}_1\cdot\boldsymbol{e}_1 + v_1w_2\boldsymbol{e}_1\cdot\boldsymbol{e}_2 + v_1w_3\boldsymbol{e}_1\cdot\boldsymbol{e}_3\\ &+ v_2w_1\boldsymbol{e}_2\cdot\boldsymbol{e}_1 + v_2w_2\boldsymbol{e}_2\cdot\boldsymbol{e}_2 + v_2w_3\boldsymbol{e}_2\cdot\boldsymbol{e}_3\\ &+ v_3w_1\boldsymbol{e}_3\cdot\boldsymbol{e}_1 + v_3w_2\boldsymbol{e}_3\cdot\boldsymbol{e}_2 + v_3w_3\boldsymbol{e}_3\cdot\boldsymbol{e}_3\,,\end{aligned}\tag{1.11}$$

which together with equation (1.8) yields the right-hand-side of (1.9). As can be seen for a very basic expansion considerable effort is required. The so called *index notation* is developed for simplification.

Index notation uses parametric index instead of explicit index. If for example the letter i is used as index in v_i it addresses any of the possible values $i = 1, 2, 3$ in three dimensional space, therefore representing any of the vector $\boldsymbol{v}$'s components. Then $\{v_i\}$ is equivalent to the set of all v_i's, namely $\{v_1, v_2, v_3\}$.

Notation 11. When an index appears twice in one term then that term is summed up over all possible values of the index. This is called *Einstein's convention.*

To clarify this, let's have a look at trace of a 3×3 matrix $\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{3} A_{ii} = A_{11} + A_{22} + A_{33}$. Based on Einstein's convention one could write $\mathrm{tr}(\boldsymbol{A}) = A_{ii}$, because the index i has appeared twice and shall be summed up over all possible values of $i = 1, 2, 3$ then $A_{ii} = A_{11} + A_{22} + A_{33}$. As another example a vector $\boldsymbol{v}$ can be written as $\boldsymbol{v} = v_i\boldsymbol{e}_i$ because the index i has appeared twice and we can write $v_i\boldsymbol{e}_i = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3$.

Remark 12. A repeating index (which is summed up over) can be freely renamed. For example $v_iw_iA_{mn}$ can be rewritten as $v_kw_kA_{mn}$ because in fact $v_iw_iA_{mn} = B_{mn}$ and index i does not appear as one of the free indices which can vary among $\{1, 2, 3\}$. Note that renaming i to m or n would not be possible as it changes the meaning.

1.7 Tensors of second order

Based on the definition of inner product, length and angle which are the building blocks of Euclidean geometry become completely general concepts (Eqs. (1.10) and (1.6)). That is why an n-dimensional vector space (Def. 7) together with an inner product operation (Prop. 10) is called a *Euclidean space* denoted by $\mathbb{E}^n$. Since we are basically unable to draw geometrical scenes in higher than three dimensions (or maybe two in fact!), it may seem just too abstract to generalize classical geometry to higher dimensions. If you remember that the physical space-time can be best expressed in a four-dimensional geometrical framework, then you might agree that this limitation of dimensions is more likely to be a restriction of our perception rather than a part of physical reality.

Another generalization would concern *projection*. Projection can be seen as geometrical transformation that can be defined based on inner product. Up to three dimensions vector projection is easily understood. Now if we allow vectors to be entities belonging to an arbitrary vector space endowed with an inner product, one can think of inner product as the means of projection of a vector in the space onto another vector or direction. Note that we have put no restrictions on the vector space such as its dimension. By the way one would ask what do we need all these cumbersome generalizations for? The answer is to understand more complex constructs based on the basic understanding we already have of components of a geometrical vector.

So far we have seen that a scalar can be expressed by a single real number. This can be seen as: scalars have no components on the three coordinate axes so require 3^0 real numbers to be specified. On the other hand a vector $\boldsymbol{v}$ has three components which are its projections onto the three coordinate axes X_i's each obtained by $v_i = \boldsymbol{v} \cdot \boldsymbol{e}_i$. Therefore, $\boldsymbol{v}$ requires 3^1 scalars to be specified.

There are more complex constructs which are called *second order tensors*. The three projections of a second order tensor $\boldsymbol{A}$ onto coordinate axes are obtained by inner product of $\boldsymbol{A}$ with basis vectors as $\boldsymbol{A}_i = \boldsymbol{A} \cdot \boldsymbol{e}_i$. These projections are vectors (not scalars!), and since each vector is determined by three scalars – its components – a second order tensor requires $3 \times 3 = 3^2$ real numbers to be specified. In array form the three components of a tensor $\boldsymbol{A}$ are vectors denoted by

$$\boldsymbol{A} = \left(\boldsymbol{A}_1 \ \boldsymbol{A}_2 \ \boldsymbol{A}_3 \right) \tag{1.12}$$

where each vector $\boldsymbol{A}_i$ has three components

$$\boldsymbol{A}_i = \begin{pmatrix} A_{1i} \\ A_{2i} \\ A_{3i} \end{pmatrix} \tag{1.13}$$

therefore $\boldsymbol{A}$ is written in the matrix form

$$\boldsymbol{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \tag{1.14}$$

where each vector $\boldsymbol{A}_i$ is written component-wise in one column of the matrix.

In analogy to the notation

$$\boldsymbol{v} = v_i \boldsymbol{e}_i \tag{1.15}$$

for vectors (based on Einstein's summation convention), a tensor can be written as

$$\boldsymbol{A} = A_{ij} \boldsymbol{e}_i \boldsymbol{e}_j \quad \text{or} \quad \boldsymbol{A} = A_{ij} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \tag{1.16}$$

with $\otimes$ being *dyadic product*. Note that we usually use the first notation for brevity.

1.8 Dyadic product

Dyadic product of two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is a tensor $\boldsymbol{A}$ such that

$$A_{ij} = v_i w_j \,. \tag{1.17}$$

Property 13. Dyadic product has the following properties

1. $\boldsymbol{v} \otimes \boldsymbol{w} \neq \boldsymbol{w} \otimes \boldsymbol{v}$
2. $\boldsymbol{u} \otimes (\alpha \boldsymbol{v} + \beta \boldsymbol{w}) = \alpha \boldsymbol{u} \otimes \boldsymbol{v} + \beta \boldsymbol{u} \otimes \boldsymbol{w}$
   $(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) \otimes \boldsymbol{w} = \alpha \boldsymbol{u} \otimes \boldsymbol{w} + \beta \boldsymbol{v} \otimes \boldsymbol{w}$
3. $(\boldsymbol{u} \otimes \boldsymbol{v}) \cdot \boldsymbol{w} = (\boldsymbol{v} \cdot \boldsymbol{w})\, \boldsymbol{u}$
   $\boldsymbol{u} \cdot (\boldsymbol{v} \otimes \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{v})\, \boldsymbol{w}$

From equation (1.16) considering the above properties we obtain

$$A_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{A} \cdot \boldsymbol{e}_j \,, \tag{1.18}$$

which is the analog of equation (1.7). In equation (1.16) the *dyads* $\boldsymbol{e}_i \boldsymbol{e}_j$ are basis tensors which are given in matrix notation as

$$\boldsymbol{e}_1\boldsymbol{e}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \boldsymbol{e}_1\boldsymbol{e}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \boldsymbol{e}_1\boldsymbol{e}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.19}$$

$$\boldsymbol{e}_2\boldsymbol{e}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \boldsymbol{e}_2\boldsymbol{e}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \boldsymbol{e}_2\boldsymbol{e}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.20}$$

$$\boldsymbol{e}_3\boldsymbol{e}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \boldsymbol{e}_3\boldsymbol{e}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \boldsymbol{e}_3\boldsymbol{e}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.21}$$

Note that **NOT** every second order tensor can be expressed as dyadic product of two vectors in general, however every second order tensor can be written as linear combination of dyadic products of vectors, as in its most common form given by the decomposition onto a given coordinate system as $\boldsymbol{A} = A_{ij} \boldsymbol{e}_i \boldsymbol{e}_j$.

1.9 Tensors of higher order

Dyadic product provides a convenient passage to definition of higher order tensors. We know that dyadic product of two vectors (first-order tensor) yields a second-order tensor. Generalization of the idea is immediate for a set of n given vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ as

$$\boldsymbol{T} = \boldsymbol{v}_1 \boldsymbol{v}_{\boldsymbol{2}} \cdots \boldsymbol{v}_n \tag{1.22}$$

which is a tensor of the n^{th} order. To make sense of it, let's take a look at inner product of the above tensor with vector $\boldsymbol{w}$

$$\boldsymbol{T} \cdot \boldsymbol{w} = (\boldsymbol{v}_1 \boldsymbol{v}_2 \cdots \boldsymbol{v}_n) \cdot \boldsymbol{w} = (\boldsymbol{v}_n \cdot \boldsymbol{w})\,(\boldsymbol{v}_1 \boldsymbol{v}_2 \cdots \boldsymbol{v}_{n-1}) \tag{1.23}$$

where $(\boldsymbol{v}_n \cdot \boldsymbol{w})$ is of course a scalar and $\boldsymbol{v}_1 \cdots \boldsymbol{v}_{n-1}$ a tensor of $(n-1)^{\text{th}}$ order. This means the tensor $\boldsymbol{T}$ maps a first-order tensor (vector) to a tensor of the $(n-1)^{\text{th}}$ order. We reemphasize that not all tensors can be written as dyadic product of vectors, however tensors can be decomposed on the basis of a coordinate system. Decomposition of the n^{th}-order tensor $\boldsymbol{T}$ is written as

$$\boldsymbol{T} = T_{i_1 i_2 \cdots i_n}\, \boldsymbol{e}_{i_1} \boldsymbol{e}_{i_2} \cdots \boldsymbol{e}_{i_n}\,. \tag{1.24}$$

where $i_1 \cdots i_n \in \{1, 2, 3\}$ and $\boldsymbol{e}_{i_1} \boldsymbol{e}_{i_2} \cdots \boldsymbol{e}_{i_n}$ is the basis tensor of the n^{th} order. It should be clear that, visualization of e.g. third-order tensors will require three-dimensional arrays and so on.

Remark 14. A vector is a *first-order tensor* and a scalar is a *zeroth-order tensor*. Order of a tensor is equal to number of its indices in index notation. Furthermore, a tensor of order n requires 3^n scalars to be specified. We usually address a second order tensor by simply calling it a tensor and the meaning should be clear from the context.

If an m^{th} order tensor $\boldsymbol{A}$ is multiplied with an n^{th} order tensor $\boldsymbol{B}$ we get an $(m+n)^{\text{th}}$ order tensor $\boldsymbol{C}$ such that

$$\begin{aligned} \boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{B} &= (A_{i_1 \cdots i_m}\, \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_m}) \otimes (B_{j_1 \cdots j_n}\, \boldsymbol{e}_{j_1} \cdots \boldsymbol{e}_{j_n}) \\ &= A_{i_1 \cdots i_m} B_{j_1 \cdots j_n} \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_m} \boldsymbol{e}_{j_1} \cdots \boldsymbol{e}_{j_n} \\ &= C_{k_1 \cdots k_{m+n}}\, \boldsymbol{e}_{k_1} \cdots \boldsymbol{e}_{k_{m+n}}\,. \end{aligned} \tag{1.25}$$

Algebra and properties of higher order tensors

On the set of n^{th}-order tensors summation operation can be defined as

$$\boldsymbol{A} + \boldsymbol{B} = (A_{i_1 \cdots i_n} + B_{i_1 \cdots i_n})\, \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_n}\,, \tag{1.26}$$

and multiplication with a scalar as

$$\alpha \boldsymbol{A} = (\alpha A_{i_1 \cdots i_n})\, \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_n}\,, \tag{1.27}$$

having the properties similar to 5 and 6. In fact tensors of order n are elements of a vector space having the dimension 3^n.

Exercise 1. Show that dimension of a vector space consisting of all n^{th}-order tensors is 3^n. (Hint: find 3^n independent tensors that can span all elements of the space.)

Generalization of inner product

Inner/Dot product of two tensors $\boldsymbol{A} = A_{i_1 \cdots i_m}\, \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_m}$ and $\boldsymbol{B} = B_{j_1 \cdots j_n}\, \boldsymbol{e}_{j_1} \cdots \boldsymbol{e}_{j_n}$ is given as

$$\begin{aligned} \boldsymbol{A} \cdot \boldsymbol{B} &= A_{i_1 \cdots i_m}\, \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_m} \cdot B_{j_1 \cdots j_n}\, \boldsymbol{e}_{j_1} \cdots \boldsymbol{e}_{j_n} \\ &= A_{i_1 \cdots i_{m-1} k} B_{k j_2 \cdots j_n} \boldsymbol{e}_{i_1} \boldsymbol{e}_{i_2} \cdots \boldsymbol{e}_{i_{m-1}} \boldsymbol{e}_{j_2} \boldsymbol{e}_{j_3} \cdots \boldsymbol{e}_{j_n} \end{aligned} \tag{1.28}$$

which is a tensor of $(m+n-2)^{\text{th}}$ order. Note that the innermost index k is common and summed up. Generalization of this idea follows

Definition 15 (contraction product)**.** For any two tensors $\boldsymbol{A}$ of the m^{th} order and $\boldsymbol{B}$ of the n^{th} order, their r-contraction product for $r \leq \min\{m,n\}$ yields a tensor $\boldsymbol{C} = \boldsymbol{A} \overset{r}{\odot} \boldsymbol{B}$ of order $(m+n-2r)$ such that

$$\boldsymbol{C} = A_{i_1 \cdots i_{m-r} k_1 \cdots k_r} \, B_{k_1 \cdots k_r j_{r+1} \cdots j_n} \, \boldsymbol{e}_{i_1} \cdots \boldsymbol{e}_{i_{m-r}} \boldsymbol{e}_{j_r} \cdots \boldsymbol{e}_{j_n} \,. \tag{1.29}$$

Note that the r innermost indices are common. The frequently used case of contraction product is the so called *double-contraction* which is denoted by $\boldsymbol{A} : \boldsymbol{B} \equiv \boldsymbol{A} \overset{2}{\odot} \boldsymbol{B}$.

Exercise 2. Write the identity $\boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\varepsilon}$ in index notation, where $\mathbb{C}$ is a fourth-order tensor and $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ are second-order tensors. Then, expand the components σ_{12} and σ_{11} for the case of two-dimensional Cartesian coordinates.

Remark 16. Contraction product of two tensors is not symmetric in general, i.e. $\boldsymbol{A} \overset{r}{\odot} \boldsymbol{B} \neq \boldsymbol{B} \overset{r}{\odot} \boldsymbol{A}$, except when $r = 1$ and $\boldsymbol{A}$ and $\boldsymbol{B}$ are first-order tensors (vectors), or $\boldsymbol{A}$ and $\boldsymbol{B}$ have certain symmetry properties.

Kronecker delta

The second order *unity tensor* or *identity tensor* denoted by $\boldsymbol{I}$ is defined by the following fundamental property.

Property 17. For every vector $\boldsymbol{v}$ it holds that

$$\boldsymbol{I} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{I} = \boldsymbol{v} \,. \tag{1.30}$$

Representation of unity tensor by index notation as

$$\delta_{ij} = \begin{cases} 1\,, & \text{if } \; i = j \\ 0\,, & \text{if } \; i \neq j\,. \end{cases} \tag{1.31}$$

is called *Kronecker's delta.* Its expansion for $i, j \in \{1,2,3\}$ gives in matrix form

$$(\delta_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \,. \tag{1.32}$$

Some important properties of Kronecker delta are

1. Orthonormality of the basis vectors $\{\boldsymbol{e}_i\}$ in equation (1.8), can be abbreviated as

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij} \,. \tag{1.33}$$

2. The so called *index exchange rule* is given as

$$u_i \delta_{ij} = u_j \quad \text{or in general} \quad A_{ij \cdots m \cdots z} \delta_{mn} = A_{ij \cdots n \cdots z} \,. \tag{1.34}$$

3. Trace of Kronecker delta

$$\delta_{ii} = 3 \,. \tag{1.35}$$

Exercise 3. Proofs are left to the reader as a practice of index notation.

Totally antisymmetric tensor

The so called *totally antisymmetric tensor* or *permutation symbol*, also named after Italian mathematician Tullio Levi-Civita as *Levi-Civita symbol*, is a third-order tensor defined as

$$\epsilon_{ijk} = \begin{cases} +1\,, & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1\,, & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0\,, & \text{otherwise, i.e. if } i=j \text{ or } j=k \text{ or } k=i. \end{cases} \tag{1.36}$$

There are important properties and relations on permutation symbol as follows

1. It is immediately known from the definition that

$$\epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji} = \epsilon_{kij} = \epsilon_{jki}\,. \tag{1.37}$$

2. Determinant of a 3×3 matrix $\boldsymbol{A}$ can be expressed by

$$\det(\boldsymbol{A}) = \epsilon_{ijk} A_{1i} A_{2j} A_{3j}\,. \tag{1.38}$$

3. The general relationship between Kronecker delta and permutation symbol reads

$$\epsilon_{ijk}\epsilon_{lmn} = \det\begin{pmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{pmatrix} \tag{1.39}$$

and as special cases

$$\epsilon_{jkl}\epsilon_{jmn} = \delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm}\,, \tag{1.40}$$

$$\epsilon_{ijk}\epsilon_{ijm} = 2\delta_{km}\,. \tag{1.41}$$

Exercise 4. Verify the identities in equations (1.40) and (1.41) based on (1.39).

1.10 Coordinate transformation

Either seen as an empirical fact or a mathematical postulate, it is well accepted that

Definition 18 (Principle of frame indifference)**.** A mathematical model of any physical phenomenon must be formulated without reference to any particular coordinate system.

This is a merit of vectors and tensors that provide mathematical models with such a generality. However in solving real world problems everything shall be expressed in terms of scalars finally for computation, and therefore introduction of a coordinate system is inevitable. The immediate question would be how to transform vectorial and tensorial quantities form one coordinate system to another. Substitution from (1.42) into the latter equation gives

$$\begin{aligned} v_1' &= (v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3)\cdot\boldsymbol{e}_1' = (\boldsymbol{e}_1'\cdot\boldsymbol{e}_1)\,v_1 + (\boldsymbol{e}_1'\cdot\boldsymbol{e}_2)\,v_2 + (\boldsymbol{e}_1'\cdot\boldsymbol{e}_3)\,v_3 \\ v_2' &= (v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3)\cdot\boldsymbol{e}_2' = (\boldsymbol{e}_2'\cdot\boldsymbol{e}_1)\,v_1 + (\boldsymbol{e}_2'\cdot\boldsymbol{e}_2)\,v_2 + (\boldsymbol{e}_2'\cdot\boldsymbol{e}_3)\,v_3 \\ v_3' &= (v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3)\cdot\boldsymbol{e}_3' = (\boldsymbol{e}_3'\cdot\boldsymbol{e}_1)\,v_1 + (\boldsymbol{e}_3'\cdot\boldsymbol{e}_2)\,v_2 + (\boldsymbol{e}_3'\cdot\boldsymbol{e}_3)\,v_3\,. \end{aligned} \tag{1.44}$$

Suppose that vector $\boldsymbol{v}$ is decomposed in Cartesian coordinates $X_1X_2X_3$ with bases $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ (Fig. 1.4)

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3 \,. \tag{1.42}$$

Since a vector is completely determined by its components, we should be able to find the decomposition of vector $\boldsymbol{v}$ on a second coordinate system $X_1'X_2'X_3'$ with bases $\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\}$ in terms of $\{v_1, v_2, v_3\}$. Using equation (1.7) we have

$$v_1' = \boldsymbol{v}\cdot\boldsymbol{e}_1' \,, \quad v_2' = \boldsymbol{v}\cdot\boldsymbol{e}_2' \,, \quad v_3' = \boldsymbol{v}\cdot\boldsymbol{e}_3' \,. \tag{1.43}$$

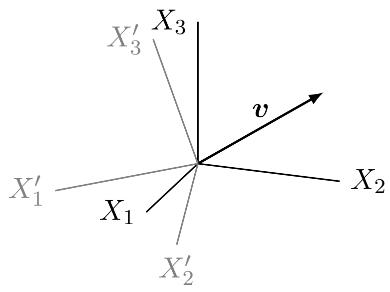

Fig. 1.4: Transformation of coordinates.

Restatement in array notation reads

$$\begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1'\cdot\boldsymbol{e}_1 & \boldsymbol{e}_1'\cdot\boldsymbol{e}_2 & \boldsymbol{e}_1'\cdot\boldsymbol{e}_3 \\ \boldsymbol{e}_2'\cdot\boldsymbol{e}_1 & \boldsymbol{e}_2'\cdot\boldsymbol{e}_2 & \boldsymbol{e}_2'\cdot\boldsymbol{e}_3 \\ \boldsymbol{e}_3'\cdot\boldsymbol{e}_1 & \boldsymbol{e}_3'\cdot\boldsymbol{e}_2 & \boldsymbol{e}_3'\cdot\boldsymbol{e}_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} , \tag{1.45}$$

and in index notation

$$v_j' = \left(\boldsymbol{e}_j'\cdot\boldsymbol{e}_i\right) v_i \,. \tag{1.46}$$

Then again, introducing *transformation tensor* $\boldsymbol{Q} = \left(\boldsymbol{e}_j'\cdot\boldsymbol{e}_i\right)$ we can rewrite the transformation rule as

$$\boldsymbol{v}' = \boldsymbol{Q}\boldsymbol{v} \,. \tag{1.47}$$

Property 19 (Orthonormal tensor)**.** Transformation $\boldsymbol{Q}$ is geometrically a rotation map with the property

$$\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I} \quad \text{or} \quad \boldsymbol{Q}^{-1} = \boldsymbol{Q}^T \tag{1.48}$$

which is called *orthonormality.*

Transformation of tensors

Having the decomposition of an n^{th}-order tensor $\boldsymbol{A}$ in $X_1X_2X_3$ coordinates as

$$\boldsymbol{A} = A_{i_1\cdots i_n}\,\boldsymbol{e}_{i_1}\cdots\boldsymbol{e}_{i_n} \,, \tag{1.49}$$

we look for its components in another coordinates $X_1'X_2'X_3'$. Following the same approach as for vectors, using equation (1.29)

$$\begin{aligned} A'_{j_1\cdots j_n} &= \boldsymbol{A} \overset{n}{\odot} \left(\boldsymbol{e}'_{j_1}\cdots\boldsymbol{e}'_{j_n}\right) \\ &= A_{i_1\cdots i_n}\,\boldsymbol{e}_{i_1}\cdots\boldsymbol{e}_{i_n} \overset{n}{\odot} \left(\boldsymbol{e}'_{j_1}\cdots\boldsymbol{e}'_{j_n}\right) \\ &= \left(\boldsymbol{e}_{i_1}\cdot\boldsymbol{e}'_{j_1}\right)\cdots\left(\boldsymbol{e}_{i_n}\cdot\boldsymbol{e}'_{j_n}\right) A_{i_1\cdots i_n} \,, \end{aligned} \tag{1.50}$$

For the special case when $\boldsymbol{A}$ is a second order tensor we can write in matrix notation

$$\boldsymbol{A}' = \boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T \,, \qquad \boldsymbol{Q} = \left(\boldsymbol{e}_j'\cdot\boldsymbol{e}_i\right) \,. \tag{1.51}$$

Remark 20. The above transformation rules are substantial. They are so fundamental that, as an alternative, we could have defined vectors and tensors based on their transformation properties instead of the geometric and algebraic approach.

Exercise 5. Find the transformation matrix for a rotation by angle θ around X_1 in three-dimensional Cartesian coordinates $X_1X_2X_3$.

Exercise 6. For a symmetric second order tensor $\boldsymbol{A}$ derive the transformation formula with a transformation given as

$$(Q_{ij}) = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

1.11 Eigenvalues and eigenvectors

The question of transformation from one coordinate system to another is answered. In practical situations we face a second question which is the choice of coordinate system in order to simplify the representation of the formulation at hand. A very basic example is explained here.

Suppose that we want to write a vector $\boldsymbol{v}$ in components. With an arbitrary choice of reference the answer is $\boldsymbol{v} = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}^T$, with $v_1, v_2, v_3 \in \mathbb{R}$. However with the choice of reference coordinates such that e.g. X_1 axis coincides with vector $\boldsymbol{v}$ we come up with $\boldsymbol{v} = \begin{pmatrix} v_1 & 0 & 0 \end{pmatrix}^T$. This second representation of the vector is simplified to one component only which is in fact a scalar, and therefore being one-dimensional. Of course in more complicated situations an appropriate choice of coordinate system can make a much greater difference. We will see that the consequences of our rather practical question turns out to be theoretically subtle. The next step in complexity is how to simplify tensors with a proper choice of coordinate system.

Definition 21. Given a second order tensor $\boldsymbol{A}$, every nonzero vector like $\boldsymbol{\psi}$ is called an *eigenvector* of $\boldsymbol{A}$ if there exists a real number λ such that

$$\boldsymbol{A} \cdot \boldsymbol{\psi} = \lambda \boldsymbol{\psi} \tag{1.52}$$

where λ is the corresponding *eigenvalue* of $\boldsymbol{\psi}$.

In general, a tensor multiplied by a vector gives another vector with a different magnitude and direction, interestingly however, there are vectors $\boldsymbol{\psi}$ on which the tensor acts like a scalar λ as laid down by equation (1.52). Now suppose that a coordinate system is chosen so that X_1 axis coincides with $\boldsymbol{\psi}$. It is clear that

$$\boldsymbol{A} \cdot \boldsymbol{e}_1 = A_{11}\boldsymbol{e}_1 + A_{21}\boldsymbol{e}_2 + A_{31}\boldsymbol{e}_3$$

and also back to (1.52)

$$\boldsymbol{A} \cdot \boldsymbol{e}_1 = \lambda \boldsymbol{e}_1 \tag{1.53}$$

comparing both equation gives

$$A_{11} = \lambda\,,\ A_{21} = 0\,,\ A_{31} = 0\,. \tag{1.54}$$

It is possible to show that for non-degenerate *symmetric* tensors in $\mathbb{R}^3$ there are three orthonormal eigenvectors which establish an orthonormal coordinate system which is called *principal coordinates*. Following equation (1.54) a symmetric tensor $\boldsymbol{\sigma}$ in principal coordinates takes the *diagonal* form

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}, \tag{1.55}$$

where its diagonal elements are eigenvalues corresponding to principal directions which are parallel to eigenvectors. The following theorem generalizes these ideas to n dimensions.

Theorem 22 (Spectral decomposition)**.** *Let $\boldsymbol{A}$ be a tensor of second order in n-dimensional space with $\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_n$ being its n linearly independent eigenvectors and their corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$. Then $\boldsymbol{A}$ can be written as*

$$\boldsymbol{A} = \boldsymbol{\Psi}\boldsymbol{\Lambda}\boldsymbol{\Psi}^{-1} \tag{1.56}$$

where $\boldsymbol{\Psi}$ is the $n \times n$ matrix having $\boldsymbol{\psi}$'s as its columns $\boldsymbol{\Psi} = (\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_n)$, and $\boldsymbol{\Lambda}$ is an $n \times n$ diagonal matrix with i^{th} diagonal element as λ_i.

Remark 23. In the case of symmetric tensors equation (1.56) takes the form of

$$\boldsymbol{A} = \boldsymbol{\Psi}\boldsymbol{\Lambda}\Psi^{T} \tag{1.57}$$

because for orthonormal eigenvectors, the matrix $\boldsymbol{\Psi}$ is an orthonormal matrix which is in fact a rotation transformation.

Remark 24. The $\boldsymbol{\Lambda}$ is the representation of $\boldsymbol{A}$ in principal coordinates. Since it is diagonal, one can write

$$\boldsymbol{\Lambda} = \sum_{i=1}^{n} \lambda_i \boldsymbol{n}_i \boldsymbol{n}_i\,. \tag{1.58}$$

Suppose that we want to derive some relationship or develop a model based on tensorial expressions. If the task is accomplished in principal coordinates it takes much less effort and at the same time, the results are completely general i.e. they hold for any other coordinate system. This is often done in material modeling and in other branches of science as well.

As the final step we would like to solve the equation (1.52). It can be recast in the form

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\,\boldsymbol{\psi} = \boldsymbol{0}\,. \tag{1.59}$$

This equation has nonzero solutions if

$$\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0 \tag{1.60}$$

which is called the *characteristic equation* of tensor $\boldsymbol{A}$. Let us have a look at its expansion

$$\begin{vmatrix} A_{11}-\lambda & A_{12} & A_{13} \\ A_{21} & A_{22}-\lambda & A_{23} \\ A_{31} & A_{32} & A_{33}-\lambda \end{vmatrix} = 0\,, \tag{1.61}$$

which gives

$$\lambda^3 - (A_{11}+A_{22}+A_{33})\,\lambda^2 + \left(\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \right) \lambda - \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} = 0 \tag{1.62}$$

which is a cubic equation with at most three distinct roots.

1.12 Invariants

Components of a vector $\boldsymbol{v}$ change from one coordinate system to another, however its length remains the same in all coordinates because length is a scalar. This is interesting because one can write the length of a vector in terms of its components in any coordinate system as $\|\boldsymbol{v}\| = (v_i v_i)^{1/2}$, which means there is a combination of components that does not change by coordinate transformation while the components themselves do. Any function of components $f(v_1, v_2, v_3)$ that does not change under coordinate transformation is called an *invariant.* This independence of coordinate system makes invariants physically important, and the reason should become clear soon.

Tensors of higher order have also invariants. For a second order tensor $\boldsymbol{A}$, looking back on equation (1.62), the values λ, λ^2 and λ^3 are scalars and stay unchanged under coordinate transformation, therefore coefficients of the equation must remain unchanged as well, and they are all invariants of tensor $\boldsymbol{A}$ denoted by

$$\mathrm{I}_{\boldsymbol{A}} = A_{11}+A_{22}+A_{33} \tag{1.63}$$

$$\mathrm{II}_{\boldsymbol{A}} = \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \tag{1.64}$$

$$\mathrm{III}_{\boldsymbol{A}} = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \tag{1.65}$$

where I, II and III are called first, second and third *principal invariants* of $\boldsymbol{A}$. The above formulas look familiar. The first invariant is trace, the second is sum of principal minors and the third is determinant. Of course, we could express invariants in terms of eigenvalues of the tensor

$$\mathrm{I}_{\boldsymbol{A}} = \mathrm{tr}(\boldsymbol{A}) = \lambda_1+\lambda_2+\lambda_3 \tag{1.66}$$

$$\mathrm{II}_{\boldsymbol{A}} = \frac{1}{2}\Big((\mathrm{tr}(\boldsymbol{A}))^2 - \mathrm{tr}\big(\boldsymbol{A}^2\big) \Big) = \lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1 \tag{1.67}$$

$$\mathrm{III}_{\boldsymbol{A}} = \det(\boldsymbol{A}) = \lambda_1\lambda_2\lambda_3\,. \tag{1.68}$$

In material modeling we often look for a scalar valued function of a tensor in the form of $\phi(\boldsymbol{A})$. Since the function value ϕ is a scalar it must be invariant under transformation of coordinates. To fulfill this requirement, the function is formulated in terms of invariants of $\boldsymbol{A}$, then it will be invariant itself. Therefore the most natural form of such a formulation would be

$$\phi = \phi(\mathrm{I}_{\boldsymbol{A}}, \mathrm{II}_{\boldsymbol{A}}, \mathrm{III}_{\boldsymbol{A}}) \ . \tag{1.69}$$

This will be used in the sequel when hyper-elastic material models are explained.

1.13 Singular value decomposition

The Euclidean norm of a vector is a non-negative scalar that represents its magnitude. It is also possible to talk about norm of a tensor. Avoiding the details, we only mention that the question of norm of a tensor $\boldsymbol{A}$ comes down to finding the maximum eigenvalue of $\boldsymbol{A}^T\boldsymbol{A}$. Since the matrix $\boldsymbol{A}^T\boldsymbol{A}$ is positive semidefinite its eigenvalues are non-negative.

Definition 25. If λ_i^2 is an eigenvalue of $\boldsymbol{A}^T\boldsymbol{A}$, then $\lambda_i > 0$ is called a singular value of $\boldsymbol{A}$.

When $\boldsymbol{A}$ is a symmetric tensor singular values of $\boldsymbol{A}$ equal the absolute values of its eigenvalues.

1.14 Cross product

Cross product of two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is a vector $\boldsymbol{u} = \boldsymbol{v} \times \boldsymbol{w}$ such that $\boldsymbol{u}$ is perpendicular to $\boldsymbol{v}$ and $\boldsymbol{w}$, and in the direction that the triple $(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{u})$ be *right-handed* (Fig. 1.5). Magnitude of $\boldsymbol{u}$ is given by

$$u = vw \sin\theta \qquad 0 \le \theta \le \pi \, . \tag{1.70}$$

The cross product of vectors has the following properties.

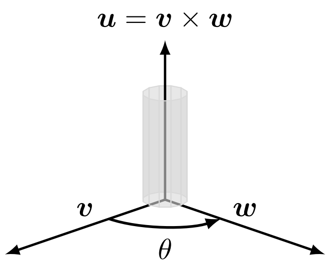

Fig. 1.5: Cross product.

Property 26 (Cross product).

1.	$\boldsymbol{v} \times \boldsymbol{w} = -\boldsymbol{w} \times \boldsymbol{v}$	Anti-symmetry
2.	$\boldsymbol{u} \times (\alpha\boldsymbol{v} + \beta\boldsymbol{w}) = \alpha\boldsymbol{u} \times \boldsymbol{v} + \beta\boldsymbol{u} \times \boldsymbol{w}$	Linearity.

Important results

1. In any orthonormal coordinate system including Cartesian coordinates it holds that

$$\begin{array}{lll} \boldsymbol{e}_1 \times \boldsymbol{e}_1 = \boldsymbol{0}\,, & \boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3\,, & \boldsymbol{e}_1 \times \boldsymbol{e}_3 = -\boldsymbol{e}_2 \\ \boldsymbol{e}_2 \times \boldsymbol{e}_1 = -\boldsymbol{e}_3\,, & \boldsymbol{e}_2 \times \boldsymbol{e}_2 = \boldsymbol{0}\,, & \boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1 \\ \boldsymbol{e}_3 \times \boldsymbol{e}_1 = \boldsymbol{e}_2\,, & \boldsymbol{e}_3 \times \boldsymbol{e}_2 = -\boldsymbol{e}_1\,, & \boldsymbol{e}_3 \times \boldsymbol{e}_3 = \boldsymbol{0}\,, \end{array} \tag{1.71}$$

2. therefore

$$\boldsymbol{v} \times \boldsymbol{w} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} , \tag{1.72}$$

3. which in index notation reads

$$\boldsymbol{v} \times \boldsymbol{w} = \epsilon_{ijk} v_j w_k \boldsymbol{e}_i \quad \text{or} \quad [\boldsymbol{v} \times \boldsymbol{w}]_i = \epsilon_{ijk} v_j w_k \,. \tag{1.73}$$

Exercise 7. Using equation (1.71) prove the identity (1.72).

Geometrical interpretation of cross product

Norm of the vector $\boldsymbol{v} \times \boldsymbol{w}$ is equal to the area of parallelogram having sides $\boldsymbol{v}$ and $\boldsymbol{w}$, and since it is normal to plane of the parallelogram we consider it to be the *area vector*. Then it is clear the the area $\boldsymbol{A}$ of a triangle with sides $\boldsymbol{v}$ and $\boldsymbol{w}$ is given by

$$\boldsymbol{A} = \frac{1}{2} \boldsymbol{v} \times \boldsymbol{w} \,. \tag{1.74}$$

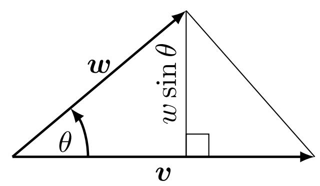

Fig. 1.6: Area of triangle.

Cross product for tensors

Cross product can be naturally extended to the case of a tensor $\boldsymbol{A}$ and a vector $\boldsymbol{v}$ as

$$\boldsymbol{A} \times \boldsymbol{v} = \epsilon_{jmn} A_{im} v_n \boldsymbol{e}_i \boldsymbol{e}_j \qquad \text{and} \qquad \boldsymbol{v} \times \boldsymbol{A} = \epsilon_{imn} v_m A_{nj} \boldsymbol{e}_i \boldsymbol{e}_j \,. \tag{1.75}$$

1.15 Triple product

For every three vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$, combination of dot and cross products gives

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) \,, \tag{1.76}$$

the so called *triple product*. Based on equation (1.72) it can be written in component form

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} , \tag{1.77}$$

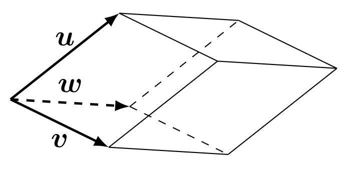

Fig. 1.7: Parallelepiped.

or in index notation

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \epsilon_{ijk} u_i v_j w_k \,. \tag{1.78}$$

Geometrically, triple product of the three vectors equals the volume of parallelepiped with sides $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ (Fig. 1.7).

1.16 Dot and cross product of vectors: miscellaneous formulas

$$\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\, \boldsymbol{v} - (\boldsymbol{u} \cdot \boldsymbol{v})\, \boldsymbol{w} \tag{1.79}$$

$$(\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{w} = (\boldsymbol{u} \cdot \boldsymbol{w})\, \boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{w})\, \boldsymbol{u} \tag{1.80}$$

$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot (\boldsymbol{w} \times \boldsymbol{x}) = (\boldsymbol{u} \cdot \boldsymbol{w})\,(\boldsymbol{v} \cdot \boldsymbol{x}) - (\boldsymbol{u} \cdot \boldsymbol{x})\,(\boldsymbol{v} \cdot \boldsymbol{w}) \tag{1.81}$$

$$\begin{aligned}(\boldsymbol{u} \times \boldsymbol{v}) \times (\boldsymbol{w} \times \boldsymbol{x}) &= (\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{x}))\, \boldsymbol{w} - (\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}))\, \boldsymbol{x} \\ &= (\boldsymbol{u} \cdot (\boldsymbol{w} \times \boldsymbol{x}))\, \boldsymbol{v} - (\boldsymbol{v} \cdot (\boldsymbol{w} \times \boldsymbol{x}))\, \boldsymbol{u} \end{aligned} \tag{1.82}$$

Exercise 8. Using index notation verify the above identities.

Exercise 9. For a tensor given as $\boldsymbol{T} = \boldsymbol{I} + \boldsymbol{v}\boldsymbol{w}$ with $\boldsymbol{v}$ and $\boldsymbol{w}$ being orthogonal vectors i.e. $\boldsymbol{v} \cdot \boldsymbol{w} = 0$, calculate $\boldsymbol{T}^2, \boldsymbol{T}^3, \cdots, \boldsymbol{T}^n$ and finally $e^{\boldsymbol{T}}$ (Hint: $e^{\boldsymbol{T}} = \sum_{n=0}^{\infty} \frac{1}{n!}\boldsymbol{T}^n$).

Chapter 2

Tensor calculus

So far, vectors and tensors are considered based on operations by which we can combine them. In this section we study vector and tensor valued functions in three dimensional space, together with their derivatives and integrals. It is important to keep in mind that we address tensors and vectors both by the term *tensor*, and the reader should already know that a vector is a first-order tensor and a scalar a zeroth-order tensor.

2.1 Tensor functions & tensor fields

A *tensor function* $\boldsymbol{A}(u)$ assigns a tensor $\boldsymbol{A}$ to a real number u. In formal mathematical notation

$$\begin{cases} \boldsymbol{A} : \mathbb{R} \longrightarrow \mathscr{T} \\ u \longmapsto \boldsymbol{A}(u) \end{cases} \tag{2.1}$$

where $\mathscr{T}$ is the space of tensors of the same order, for instance second order tensors $\mathscr{T} = \mathbb{R}^{3\times3}$ or vectors $\mathscr{T} = \mathbb{R}^3$ or maybe fourth order tensors $\mathscr{T} = \mathbb{R}^{3\times3\times3\times3}$.

A typical example is the position vector of a moving particle given as a function of time $\boldsymbol{x}(t)$ where t is a real number and $\boldsymbol{x}$ is a vector with initial point at origin and terminal point at the moving particle (Fig. 2.1).

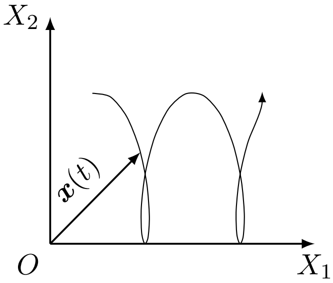

Fig. 2.1: Position vector.

On the other hand a *tensor filed* $\boldsymbol{A}(\boldsymbol{x})$ assigns a tensor $\boldsymbol{A}$ to a point $\boldsymbol{x}$ in space, say three dimensional. Again one could formally write

$$\begin{cases} \boldsymbol{A} : \mathbb{R}^3 \longrightarrow \mathscr{T} \\ \boldsymbol{x} \longmapsto \boldsymbol{A}(\boldsymbol{x}) \end{cases} . \tag{2.2}$$

Here, $\mathbb{R}^3$ is the usual three dimensional Euclidean space. For instance, we can think of temperature distribution in a given domain $T(\boldsymbol{x})$ which is a scalar field that assigns scalars (temperatures) to points $\boldsymbol{x}$ in the domain. Another example could be strain filed $\boldsymbol{\varepsilon}(\boldsymbol{x})$ in a solid body under deformation, which assigns second-order tensors $\boldsymbol{\varepsilon}$ to points $\boldsymbol{x}$ in the body.

2.2 Derivative of tensor functions

Derivative of a differentiable tensor function $\boldsymbol{A}(u)$ is defined as

$$\frac{d\boldsymbol{A}}{du} = \lim_{\Delta u \to 0} \frac{\boldsymbol{A}(u + \Delta u) - \boldsymbol{A}(u)}{\Delta u}. \tag{2.3}$$

We assume that all derivatives exist unless otherwise stated. For the case of a vector function $\boldsymbol{A}(u) = A_i(u)\,\boldsymbol{e}_i$ in Cartesian coordinates the above definition becomes

$$\frac{d\boldsymbol{A}}{du} = \frac{dA_i}{du}\boldsymbol{e}_i = \frac{dA_1}{du}\boldsymbol{e}_1 + \frac{dA_2}{du}\boldsymbol{e}_2 + \frac{dA_3}{du}\boldsymbol{e}_3. \tag{2.4}$$

Because $\boldsymbol{e}_i$'s remain unchanged in Cartesian coordinates their derivatives are zero.

Exercise 10. For a given second-order tensor function $\boldsymbol{A}(u) = A_{ij}(u)\,\boldsymbol{e}_i\boldsymbol{e}_j$ in Cartesian coordinates write down the derivatives.

Property 27. For tensor functions $\boldsymbol{A}(u), \boldsymbol{B}(u), \boldsymbol{C}(u)$, and vector function $\boldsymbol{a}(u)$ we have

1. $$\frac{d}{du}(\boldsymbol{A}\cdot\boldsymbol{B}) = \boldsymbol{A}\cdot\frac{d\boldsymbol{B}}{du} + \frac{d\boldsymbol{A}}{du}\cdot\boldsymbol{B} \tag{2.5}$$
2. $$\frac{d}{du}(\boldsymbol{A}\times\boldsymbol{B}) = \boldsymbol{A}\times\frac{d\boldsymbol{B}}{du} + \frac{d\boldsymbol{A}}{du}\times\boldsymbol{B} \tag{2.6}$$
3. $$\frac{d}{du}[\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})] = \frac{d\boldsymbol{A}}{du}\cdot(\boldsymbol{B}\times\boldsymbol{C}) + \boldsymbol{A}\cdot\left(\frac{d\boldsymbol{B}}{du}\times\boldsymbol{C}\right) + \boldsymbol{A}\cdot\left(\boldsymbol{B}\times\frac{d\boldsymbol{C}}{du}\right) \tag{2.7}$$
4. $$\boldsymbol{a}\cdot\frac{d\boldsymbol{a}}{du} = a\frac{da}{du} \tag{2.8}$$
5. $$\boldsymbol{a}\cdot\frac{d\boldsymbol{a}}{du} = 0 \qquad \text{iff} \qquad \|\boldsymbol{a}\| = \texttt{const} \tag{2.9}$$

Exercise 11. Verify the 4^{th} and the 5^{th} properties above.

2.3 Derivatives of tensor fields, Gradient, Nabla operator

Since a tensor filed $\boldsymbol{A}(\boldsymbol{x})$ is a multi-variable function its differential $d\boldsymbol{A}$ depends on the direction of differential of independent variable $d\boldsymbol{x}$ which is a vector $d\boldsymbol{x} = dx_i\boldsymbol{e}_i$. When $d\boldsymbol{x}$ is parallel to one of coordinate axes, partial derivatives of the tensor field $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}(x_1, x_2, x_3)$ are obtained

$$\begin{aligned}
\frac{\partial\boldsymbol{A}}{\partial x_1} &= \lim_{\Delta x_1 \to 0} \frac{\boldsymbol{A}(x_1 + \Delta x_1, x_2, x_3) - \boldsymbol{A}(x_1, x_2, x_3)}{\Delta x_1} \\
\frac{\partial\boldsymbol{A}}{\partial x_2} &= \lim_{\Delta x_2 \to 0} \frac{\boldsymbol{A}(x_1, x_2 + \Delta x_2, x_3) - \boldsymbol{A}(x_1, x_2, x_3)}{\Delta x_2} \\
\frac{\partial\boldsymbol{A}}{\partial x_3} &= \lim_{\Delta x_3 \to 0} \frac{\boldsymbol{A}(x_1, x_2, x_3 + \Delta x_3) - \boldsymbol{A}(x_1, x_2, x_3)}{\Delta x_3}
\end{aligned} \tag{2.10}$$

or in compact form

$$\frac{\partial \boldsymbol{A}}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{\boldsymbol{A}(\boldsymbol{x} + \Delta x_i \boldsymbol{e}_i) - \boldsymbol{A}(\boldsymbol{x})}{\Delta x_i} . \tag{2.11}$$

In general, $d\boldsymbol{x}$ is not parallel to coordinate axes. Therefore based on the *chain rule* of differentiation

$$d\boldsymbol{A}(x_1, x_2, x_3) = \frac{\partial \boldsymbol{A}}{\partial x_i} dx_i = \frac{\partial \boldsymbol{A}}{\partial x_1} dx_1 + \frac{\partial \boldsymbol{A}}{\partial x_2} dx_3 + \frac{\partial \boldsymbol{A}}{\partial x_3} dx_3 , \tag{2.12}$$

Remembering from calculus of multi-variable functions, this can also be written as

$$d\boldsymbol{A} = \operatorname{grad}(\boldsymbol{A}) \cdot d\boldsymbol{x} \quad \text{with} \quad \operatorname{grad}(\boldsymbol{A}) = \frac{\partial \boldsymbol{A}}{\partial x_i} \boldsymbol{e}_i . \tag{2.13}$$

In the latter formula *gradient* of tensor field grad($\boldsymbol{A}$) appears which is a tensor of one order higher than $\boldsymbol{A}$ itself. To highlight this fact, we rewrite the above equation as

$$\operatorname{grad}(\boldsymbol{A}) = \frac{\partial \boldsymbol{A}}{\partial x_i} \otimes \boldsymbol{e}_i . \tag{2.14}$$

A proper choice of notation in mathematics is substantial and introduction of operators notation in calculus is no exception. Since linear operators follow algebraic rules somewhat similar to that of arithmetics of real numbers, writing relations in operators notation is invaluable. Here, we introduce the so called *nabla operator* as

$$\nabla \equiv \boldsymbol{e}_i \frac{\partial}{\partial x_i} = \boldsymbol{e}_1 \frac{\partial}{\partial x_1} + \boldsymbol{e}_2 \frac{\partial}{\partial x_2} + \boldsymbol{e}_3 \frac{\partial}{\partial x_3} . \tag{2.15}$$

Then gradient of the tensor filed in equation (2.13) can be written as*

$$\operatorname{grad}(\boldsymbol{A}) = \boldsymbol{A}\nabla = \boldsymbol{A} \otimes \nabla . \tag{2.16}$$

2.4 Integrals of tensor functions

If $\boldsymbol{A}(u)$ and $\boldsymbol{B}(u)$ are tensor functions such that $\boldsymbol{A}(u) = {}^{d}\!/_{du}\, \boldsymbol{B}(u)$, the *indefinite* integral of $\boldsymbol{A}$ is

$$\int \boldsymbol{A}(u)\; du = \boldsymbol{B}(u) + \boldsymbol{C} \tag{2.17}$$

with $\boldsymbol{C} = \texttt{const}$, and the *definite* integral of $\boldsymbol{A}$ over the interval $[a, b]$ is

$$\int_a^b \boldsymbol{A}(u)\; du = \boldsymbol{B}(b) - \boldsymbol{B}(a)\; . \tag{2.18}$$

Decomposition of equation (2.17) for (say) a second-order tensor gives

$$\int A_{ij}(u)\; du = B_{ij}(u) + C_{ij} , \tag{2.19}$$

where A_{ij} is a real valued function of real numbers. Therefore, integration of a tensor function comes down to integration of its components.

*In recent equations the dyadic $\otimes$ is written explicitly sometimes for pedagogical reasons.

2.5 Line integrals

Consider a space curve C joining two points $P_1\,(a_1, a_2, a_3)$ and $P_2\,(b_1, b_2, b_3)$ as shown in Fig. 2.2. If the curve is divided into N parts by subdivision points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}$, then the line integral of a tensor field $\boldsymbol{A}(\boldsymbol{x})$ along C is given by

$$\int_C \boldsymbol{A} \cdot d\boldsymbol{x} = \int_{P_1}^{P_2} \boldsymbol{A} \cdot d\boldsymbol{x} = \lim_{N\to\infty} \sum_{i=1}^{N} \boldsymbol{A}(\boldsymbol{x}_i) \cdot \Delta \boldsymbol{x}_i \qquad (2.20)$$

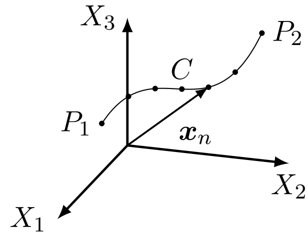

Fig. 2.2: Space curve.

where $\Delta \boldsymbol{x}_i = \boldsymbol{x}_i - \boldsymbol{x}_{i-1}$, and when $N \to \infty$ the largest of divisions tends to zero

$$\max_i \|\Delta \boldsymbol{x}_i\| \to 0\,.$$

In Cartesian coordinates the line integral can be written as

$$\int_C \boldsymbol{A} \cdot d\boldsymbol{x} = \int_C (A_1\, dx_1 + A_2\, dx_2 + A_3\, dx_3)\,. \qquad (2.21)$$

Line integral has the following properties.

Property 28.

1. $$\int_{P_1}^{P_2} \boldsymbol{A} \cdot d\boldsymbol{x} = -\int_{P_2}^{P_1} \boldsymbol{A} \cdot d\boldsymbol{x} \qquad (2.22)$$
2. $$\int_{P_1}^{P_2} \boldsymbol{A} \cdot d\boldsymbol{x} = \int_{P_1}^{P_3} \boldsymbol{A} \cdot d\boldsymbol{x} + \int_{P_3}^{P_2} \boldsymbol{A} \cdot d\boldsymbol{x} \qquad P_3 \text{ is between } P_1 \text{ and } P_2 \qquad (2.23)$$

Parameterization

A curve in space is a one-dimensional entity which can be specified by a parameterization of the form

$$\boldsymbol{x}(s) \qquad (2.24)$$

which obviously means

$$x_1 = x_1(s)\ ,\ x_2 = x_2(s)\ ,\ x_3 = x_3(s)\ .$$

Therefore a path integral can be parametrized in terms of s by

$$\int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \boldsymbol{A} \cdot d\boldsymbol{x} = \int_{s_1}^{s_2} \boldsymbol{A} \cdot \frac{d\boldsymbol{x}}{ds}\, ds\,. \qquad (2.25)$$

2.6 Path independence

A line integral as introduced above, depends on its limit points P_1 and P_2, on particular choice of the path C, and on its integrand $\boldsymbol{A}(\boldsymbol{x})$. However, there are cases where the line integral is independent of the path C. Then, having specified the tensor field $\boldsymbol{A}(\boldsymbol{x})$ and the initial point P_1, the line integral is a function of the position of P_2 only. If we denote the position of P_2 by $\boldsymbol{x}$ then

$$\boldsymbol{\phi}(\boldsymbol{x}) = \int_{P_1}^{\boldsymbol{x}} \boldsymbol{A} \cdot d\boldsymbol{x} \,. \tag{2.26}$$

Now if we consider differentiating the tensor field $\boldsymbol{\phi}(\boldsymbol{x})$ it yields

$$d\boldsymbol{\phi} = \boldsymbol{A} \cdot d\boldsymbol{x} \,, \tag{2.27}$$

which according definition of gradient (2.13) means

$$\boldsymbol{A} = \mathrm{grad}(\boldsymbol{\phi}) = \boldsymbol{\phi}\nabla \,. \tag{2.28}$$

This argument could be reversed which would lead to the following theorem.

Theorem 29. *The line integral $\int_C \boldsymbol{A} \cdot d\boldsymbol{x}$ is path independent if and only if there is a function $\boldsymbol{\phi}(\boldsymbol{x})$ such that $\boldsymbol{A} = \boldsymbol{\phi}\nabla$.*

Exercise 12. Prove the reverse argument of the above theorem. That is, starting from existence of $\boldsymbol{\phi}(\boldsymbol{x})$ such that $\boldsymbol{A} = \boldsymbol{\phi}\nabla$, prove that the line integral is path independent.

Remark 30. Note that in particular case of C being a *closed path* the line integral is path independent if and only if

$$\oint_C \boldsymbol{A} \cdot d\boldsymbol{x} = \boldsymbol{0} \,, \tag{2.29}$$

for arbitrary path C. The circle on integral sign shows that C is closed. This is another statement of path-independence which is equivalent to the ones mentioned before.

2.7 Surface integrals

Consider a smooth connected surface S in space (Fig. 2.3) subdivided into N elements ΔS_i, $i = 1, \ldots, N$ at positions $\boldsymbol{x}_i$ with unit normal vectors to the surface $\boldsymbol{n}_i$. Let tensor field $\boldsymbol{A}(\boldsymbol{x})$ be defined over the domain $\Omega \supseteq S$. Then the surface integral of *normal* component of $\boldsymbol{A}$ over S is defined as

$$\int_S \boldsymbol{A} \cdot \boldsymbol{n} \, dS = \lim_{N\to\infty} \sum_{i=1}^{N} \boldsymbol{A}(\boldsymbol{x}_i) \cdot \boldsymbol{n}_i \, \Delta S_i \tag{2.30}$$

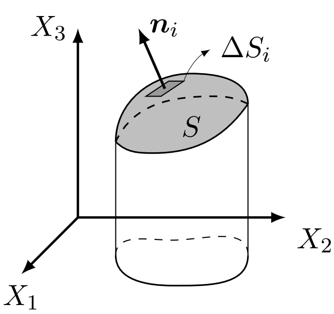

Fig. 2.3: Surface integral.

where it holds that

$$N \to \infty \quad : \quad \max_i \|\Delta S_i\| \to 0 \,.$$

The surface integral of *tangential* component of $\boldsymbol{A}$ over S is defined as

$$\int_S \boldsymbol{A} \times \boldsymbol{n}\, dS = \lim_{N\to\infty} \sum_{i=1}^{N} \boldsymbol{A}(\boldsymbol{x}_i) \times \boldsymbol{n}_i\, \Delta S_i \tag{2.31}$$

Note the surface integral (2.30) is a tensor of one order lower than $\boldsymbol{A}$ due to appearance of $\boldsymbol{A}\cdot\boldsymbol{n}$ in the integrand, while the surface integral (2.31) is a tensor of the same order as $\boldsymbol{A}$ due to the term $\boldsymbol{A}\times\boldsymbol{n}$.

Surface integrals in terms of double-integrals

Surface integrals can be calculated in Cartesian coordinates. This is the case usually when the surface S does not have any symmetries, such as rotational symmetry. For this the surface S is projected onto Cartesian plane X_1X_2 which gives a domain like $\mathcal{S}$ (Fig. 2.3). Then

$$\int_S \boldsymbol{A}\cdot\boldsymbol{n}\, dS = \iint\limits_{\mathcal{S}} \boldsymbol{A}\cdot\boldsymbol{n}\,\frac{dx_1\, dx_2}{\boldsymbol{n}\cdot\boldsymbol{e}_3}\,. \tag{2.32}$$

The transformation between the area element dS and its projection dx_1dx_2 is given by

$$dx_1dx_2 = (\boldsymbol{n}dS)\cdot\boldsymbol{e}_3 = (\boldsymbol{n}\cdot\boldsymbol{e}_3)\, dS\,,$$

where $\boldsymbol{n}$ is the unit normal vector to dS and $\boldsymbol{e}_3$ is the unit normal vector to dx_1dx_2.

Exercise 13. Calculate the surface integral of the normal component of vector field $\boldsymbol{v} = {}^1\!/_{r^2}\, \boldsymbol{e}_r$ over the unit sphere centered at origin.

2.8 Divergence and curl operators

In surface integrals introduced above, the case when the surface S is closed has a special importance. In a physical context, integral of the normal component of a tensor field over a closed surface expresses the overall *flux* of the field through the surface which is a measure of intensity of *sources*†. On the other hand, closed surface integral of the tangential component represents the so called *rotation* of the field around the surface.

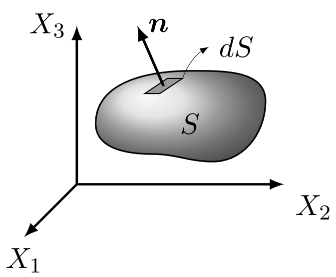

Fig. 2.4: Closed surface.

Assume that S is a closed surface with *outward* unit normal vector $\boldsymbol{n}$, occupying the domain Ω with volume V. Integral over closed surfaces is usually denoted by $\oint$. We can define surface integrals of a given tensor field $\boldsymbol{A}(\boldsymbol{x})$ over S as in equations (2.30) and (2.31). As mentioned before, in the case of closed surfaces these have special physical meanings. A question that naturally arises is: what are the average intensity of sources and rotations over the domain Ω. The answer is

$$\mathcal{D} = \frac{1}{V}\oint_S \boldsymbol{A}\cdot\boldsymbol{n}\, dS \qquad \mathcal{C} = \frac{1}{V}\oint_S \boldsymbol{A}\times\boldsymbol{n}\, dS\,, \tag{2.33}$$

†A source, as it literally suggests, is what causes the tensor field.

where integrals are divided by volume V. Now if the domain Ω shrinks, that is $V \to 0$, then these average values reflect the local intensity or *concentration* of sources and rotations of the tensor field. This is the physical analogue of density which is the concentration of mass. These local values are called *divergence* and *curl* of the tensor field $\boldsymbol{A}$, denoted and formally defined by

$$\operatorname{div}(\boldsymbol{A}) = \lim_{V\to 0} \frac{1}{V} \oint_S \boldsymbol{A} \cdot \boldsymbol{n}\, dS \tag{2.34}$$

$$\operatorname{curl}(\boldsymbol{A}) = \lim_{V\to 0} \frac{1}{V} \oint_S \boldsymbol{A} \times \boldsymbol{n}\, dS\,. \tag{2.35}$$

It can be shown that

$$\operatorname{div}(\boldsymbol{A}) = \boldsymbol{A} \cdot \nabla \tag{2.36}$$

$$\operatorname{curl}(\boldsymbol{A}) = \boldsymbol{A} \times \nabla \tag{2.37}$$

in operator notation. The proof is straightforward in Cartesian coordinates and can be found in most calculus books.

Exercise 14. Using a good calculus book, starting from (2.34) and (2.35) derive equations (2.36) and (2.37) in Cartesian coordinates for a vector field $\boldsymbol{u}(\boldsymbol{x})$ as

$$\operatorname{div}(\boldsymbol{u}) = \boldsymbol{u} \cdot \nabla = u_i \partial_i \quad \text{and} \quad \operatorname{curl}(\boldsymbol{u}) = \boldsymbol{u} \times \nabla = \epsilon_{kij} u_i \partial_j\,.$$

Remark 31. So far gradient, divergence and curl operators are applied to their operand form the right-hand-side i.e. $\operatorname{grad}(\boldsymbol{A}) = \boldsymbol{A}\nabla$, $\operatorname{div}(\boldsymbol{A}) = \boldsymbol{A} \cdot \nabla$, $\operatorname{curl}(\boldsymbol{A}) = \boldsymbol{A} \times \nabla$. This is not a strict requirement by their definition. In fact the direction from which the operator is applied depends on the context. If we exchange the order of dyadic, inner and outer products in equations (2.26), (2.34) and (2.35) respectively, the direction of application of nabla will be reversed, to wit

$$\boldsymbol{\phi} = \int d\boldsymbol{x} \cdot \nabla\boldsymbol{\phi} \qquad \lim_{V\to 0} \frac{1}{V} \oint_S \boldsymbol{n} \cdot \boldsymbol{A}\, dS = \nabla \cdot \boldsymbol{A} \qquad \lim_{V\to 0} \frac{1}{V} \oint_S \boldsymbol{n} \times \boldsymbol{A}\, dS = \nabla \times \boldsymbol{A} \tag{2.38}$$

which are equally valid. However, keeping a consistence convention shall not be overlooked.

Exercise 15. Calculate $\nabla \times (\nabla \times \boldsymbol{v})$ for the vector field $\boldsymbol{v}$.

2.9 Laplace's operator

We remember from equation (2.28) that a tensor field $\boldsymbol{A}$ whose line integrals are path-independent can be written as the gradient of a lower-order tensor field as $\boldsymbol{A} = \boldsymbol{\phi}\nabla$. Now if we are interested in the concentration of sources that generate $\boldsymbol{A}$ then

$$\operatorname{div}(\boldsymbol{A}) = \operatorname{div}(\operatorname{grad}(\boldsymbol{\phi})) = \boldsymbol{\phi}\nabla \cdot \nabla \tag{2.39}$$

This combination of gradient and divergence operators appears so often that it is given a distinct name, the so called *Laplace's operator* or *Laplacian* which is denoted by

$$\Delta \equiv \nabla \cdot \nabla = \partial_k \partial_k\,. \tag{2.40}$$

Laplacian of a field $\Delta \boldsymbol{A}$ is sometimes denoted by $\nabla^2 \boldsymbol{A}$. Laplace's operator is also called *harmonic* operator. That is why two consecutive application of Laplacian over a field is called *biharmonic* operator denoted by

$$\nabla^4 \boldsymbol{A} = \Delta\Delta \boldsymbol{A} = \Delta\left(\Delta \boldsymbol{A}\right) \tag{2.41}$$

which is expanded in Cartesian coordinates as

$$\Delta\Delta\phi = \frac{\partial^4 \phi}{\partial x_1^4} + \frac{\partial^4 \phi}{\partial x_2^4} + \frac{\partial^4 \phi}{\partial x_3^4} + 2\frac{\partial^4 \phi}{\partial x_1^2 \partial x_2^2} + 2\frac{\partial^4 \phi}{\partial x_2^2 \partial x_3^2} + 2\frac{\partial^4 \phi}{\partial x_3^2 \partial x_1^2}\,. \tag{2.42}$$

2.10 Gauss' and Stoke's Theorems

Theorem 32 (Gauss theorem)**.** *Let S be a closed surface bounding a region Ω of volume V with outward unit normal vector $\boldsymbol{n}$, and $\boldsymbol{A}(\boldsymbol{x})$ a tensor field (Fig. 2.4). Then we have*

$$\int_V \boldsymbol{A} \cdot \nabla \, dV = \oint_S \boldsymbol{A} \cdot \boldsymbol{n} \, dS\,. \tag{2.43}$$

This result is also called *Green's* or *divergence* theorem. If we remember the physical interpretation of divergence, then the Gauss theorem means the flux of a tensor field through a closed surface equals the intensity of the sources bounded by the surface, which physically makes perfect sense.

Theorem 33 (Stoke's theorem)**.** *Let S be an open two-sided surface bounded by a closed non-self-intersecting curve C as in Fig. 2.5. Then*

$$\oint_C \boldsymbol{A} \cdot d\boldsymbol{x} = \int_S (\boldsymbol{A} \times \nabla) \cdot \boldsymbol{n} \, dS \tag{2.44}$$

where $\boldsymbol{n}$ is directed towards counter-clockwise *view of the path direction as indicated in the figure.*

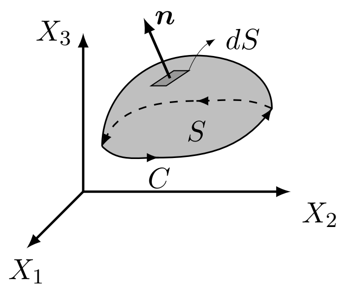

Fig. 2.5: Stoke's theorem.

Physical interpretation of Stoke's theorem is much similar to that of Gauss' theorem except that it involves closed line integration and enclosed surface, instead of closed surface integration and enclosed volume. Namely, the overall rotation of the tensor field along a closed path equals the intensity of rotations enclosed by the path.

2.11 Miscellaneous formulas and theorems involving nabla

For given vector fields $\boldsymbol{v}$ and $\boldsymbol{w}$, tensor fields $\boldsymbol{A}$ and $\boldsymbol{B}$, and scalar fields α and β we have

$$\nabla(\boldsymbol{A}+\boldsymbol{B}) = \nabla\boldsymbol{A} + \nabla\boldsymbol{B} \tag{2.45}$$
$$\nabla\cdot(\boldsymbol{A}+\boldsymbol{B}) = \nabla\cdot\boldsymbol{A} + \nabla\cdot\boldsymbol{B} \tag{2.46}$$
$$\nabla\times(\boldsymbol{A}+\boldsymbol{B}) = \nabla\times\boldsymbol{A} + \nabla\times\boldsymbol{B} \tag{2.47}$$
$$\nabla\cdot(\alpha\boldsymbol{A}) = (\nabla\alpha)\cdot\boldsymbol{A} + \alpha(\nabla\cdot\boldsymbol{A}) \tag{2.48}$$
$$\nabla\times(\alpha\boldsymbol{A}) = (\nabla\alpha)\times\boldsymbol{A} + \alpha(\nabla\times\boldsymbol{A}) \tag{2.49}$$
$$\nabla\times(\nabla\times\boldsymbol{A}) = \nabla(\nabla\cdot\boldsymbol{A}) - \nabla^2\boldsymbol{A} \tag{2.50}$$
$$\nabla\cdot(\boldsymbol{v}\times\boldsymbol{w}) = (\nabla\times\boldsymbol{v})\cdot\boldsymbol{w} + \boldsymbol{v}\cdot(\nabla\times\boldsymbol{w}) \tag{2.51}$$
$$\nabla\times(\boldsymbol{v}\times\boldsymbol{w}) = (\boldsymbol{w}\cdot\nabla)\boldsymbol{v} - (\nabla\cdot\boldsymbol{v})\boldsymbol{w} - (\boldsymbol{v}\cdot\nabla)\boldsymbol{w} + (\nabla\cdot\boldsymbol{w})\boldsymbol{v} \tag{2.52}$$
$$\nabla(\boldsymbol{v}\cdot\boldsymbol{w}) = (\boldsymbol{w}\cdot\nabla)\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{w} + \boldsymbol{w}\times(\nabla\times\boldsymbol{v}) + \boldsymbol{v}\times(\nabla\times\boldsymbol{w}) \tag{2.53}$$

Theorem 34. *For a tensor field* $\boldsymbol{A}$

$$\nabla\times(\nabla\boldsymbol{A}) = \boldsymbol{0} \qquad (\boldsymbol{A}\nabla)\times\nabla = \boldsymbol{0} \tag{2.54}$$
$$\nabla\cdot(\nabla\times\boldsymbol{A}) = \boldsymbol{0} \qquad (\boldsymbol{A}\times\nabla)\cdot\nabla = \boldsymbol{0} \tag{2.55}$$

given that $\boldsymbol{A}$ *is smooth enough.*

Let us rephrase the above theorem as follows:

- Gradient of a tensor field is *curl-free*, i.e. its curl vanishes.
- Curl of a tensor field is *divergence-free*, i.e. its divergence vanishes.

Comparing to Gauss and Stokes's theorems, these mean

$$\oint_S (\boldsymbol{A}\times\nabla)\cdot\boldsymbol{n}\,dS = \int_V (\boldsymbol{A}\times\nabla)\cdot\nabla\,dV = \boldsymbol{0} \tag{2.56}$$

and

$$\oint_C (\boldsymbol{A}\nabla)\cdot d\boldsymbol{x} = \int_S (\boldsymbol{A}\nabla)\times\nabla\cdot\boldsymbol{n}\,dS = \boldsymbol{0}\,. \tag{2.57}$$

Exercise 16. Show that in general $\nabla\cdot\boldsymbol{A}\times\nabla \neq \boldsymbol{0}$ and $\nabla\times\boldsymbol{A}\cdot\nabla \neq \boldsymbol{0}$.

Remark 35. The recent equation, according (2.29), is equivalent to path-independence.

Theorem 36 (Green's identities)**.** *For two scalar fields* ϕ *and* ψ *the first Green's identity is*

$$\int_V \left[\phi\,(\psi\Delta) + (\phi\nabla)\cdot(\psi\nabla)\right]\,dV = \oint_S \phi\,(\psi\nabla)\cdot\boldsymbol{n}\,dS \tag{2.58}$$

and the second Green's identity

$$\int_V \left[\phi\,(\psi\Delta) - (\phi\Delta)\,\psi\right]\,dV = \oint_S \left[\phi\,(\psi\nabla) - (\phi\nabla)\,\psi\right]\cdot\boldsymbol{n}\,dS\,. \tag{2.59}$$

Alternative forms of Gauss' and Stoke's theorems

For tensor field $\boldsymbol{A}$ and scalar field ϕ we have

$$\int_V \boldsymbol{A} \times \nabla \, dV = \oint_S \boldsymbol{A} \times \boldsymbol{n} \, dS \qquad \oint_C \phi \, d\boldsymbol{x} = \int_S (\phi \nabla) \times \boldsymbol{n} \, dS \, . \tag{2.60}$$

which could be equivalently stated as

$$\int_V \nabla \times \boldsymbol{A} \, dV = \oint_S \boldsymbol{n} \times \boldsymbol{A} \, dS \qquad \oint_C \phi \, d\boldsymbol{x} = \int_S \boldsymbol{n} \times (\nabla \phi) \, dS \, . \tag{2.61}$$

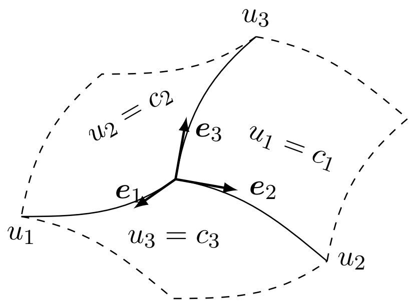

Fig. 2.6: Curvilinear coordinates.

2.12 Orthogonal curvilinear coordinate systems

A point P in space (Fig. 2.6) can be specified by its rectangular coordinates (another name for Cartesian coordinates) (x_1, x_2, x_3), or curvilinear coordinates (u_1, u_2, u_3). Then there must be transformation rules of the form

$$\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{u}) \quad \text{and} \quad \boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}) \tag{2.62}$$

where the transformation functions are *smooth* and *on-to-one*. Smoothness means being differentiable, and being one-to-one requires *Jacobian* determinant to be non-zero

$$\det\left(\frac{\partial (u_1, u_2, u_3)}{\partial (x_1, x_2, x_3)}\right) = \begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \frac{\partial u_1}{\partial x_3} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \frac{\partial u_2}{\partial x_3} \\ \frac{\partial u_3}{\partial x_1} & \frac{\partial u_3}{\partial x_2} & \frac{\partial u_3}{\partial x_3} \end{vmatrix} \neq 0 \tag{2.63}$$

Definition 37 (Jacobian)**.** For a differentiable function $\boldsymbol{f}(\boldsymbol{x})$ such that $\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^n$, its Jacobian is an $m \times n$ matrix defined as $\boldsymbol{J} = \left(\frac{\partial f_i}{\partial x_j}\right)$ with $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

At the point P there are three surfaces described by

$$u_1 = \texttt{const} \, , \quad u_2 = \texttt{const} \, , \quad u_3 = \texttt{const} \tag{2.64}$$

passing through P, called *coordinate surfaces*. And there are also three curves which are intersections of every two coordinate surfaces. These are called *coordinate curves*. On each coordinate curve only one of the three coordinates u_i varies and the other two are constant. If the position vector is denoted by $\boldsymbol{r}$ then the vector $\partial \boldsymbol{r}/\partial u_i$ is tangent to u_i coordinate curve, and the unit tangent vectors (Fig. 2.6) are obtained from

$$\boldsymbol{e}_1 = \frac{\partial \boldsymbol{r}/\partial u_1}{\|\partial \boldsymbol{r}/\partial u_1\|}\,, \quad \boldsymbol{e}_2 = \frac{\partial \boldsymbol{r}/\partial u_2}{\|\partial \boldsymbol{r}/\partial u_2\|}\,, \quad \boldsymbol{e}_3 = \frac{\partial \boldsymbol{r}/\partial u_3}{\|\partial \boldsymbol{r}/\partial u_3\|} \tag{2.65}$$

Introducing scale factors $h_i = \|\partial \boldsymbol{r}/\partial u_i\|$ we have

$$\partial \boldsymbol{r}/\partial u_1 = h_1 \boldsymbol{e}_1\,, \quad \partial \boldsymbol{r}/\partial u_2 = h_2 \boldsymbol{e}_2\,, \quad \partial \boldsymbol{r}/\partial u_3 = h_3 \boldsymbol{e}_3 \tag{2.66}$$

If $\boldsymbol{e}_1, \boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are mutually orthogonal then we call (u_1, u_2, u_3) an *orthogonal curvilinear coordinate system.*

Remark 38. In rectangular coordinates the coordinate surfaces are flat planes and coordinate curves are straight lines. Also, the scale factors h_i's are equal to one due to straight coordinate curves. Furthermore, the unit vectors $\boldsymbol{e}_i$'s are the same for all points in space. These properties do not generally hold for curvilinear coordinate systems.

2.13 Differentials

In three dimensional space, there are three types of geometrical objects (other than points) regarding dimensions

- one-dimensional objects: lines
- two-dimensional objects: surfaces
- three-dimensional objects: volumes.

Each type has its own differential element, namely line elements, surface elements and volume elements. We assume that the reader has already dealt with the related ideas in rectangular (Cartesian) coordinates. Now we want to generalize those ideas to curvilinear coordinates.

Line elements

An infinitesimal line segment with an arbitrary direction in space can be projected onto coordinate axes. Consider the position vector $\boldsymbol{r}$ explained in a curvilinear coordinate system by $\boldsymbol{r}(u_1, u_2, u_3)$. A line element is the differential of the position vector obtained by chain rule and using equation (2.66)

$$d\boldsymbol{r} = \frac{\partial \boldsymbol{r}}{\partial u_1} du_1 + \frac{\partial \boldsymbol{r}}{\partial u_2} du_2 + \frac{\partial \boldsymbol{r}}{\partial u_3} du_3 = h_1\, du_1\, \boldsymbol{e}_1 + h_2\, du_2\, \boldsymbol{e}_2 + h_3\, du_3\, \boldsymbol{e}_3\,. \tag{2.67}$$

where

$$dr_1 = h_1\, du_1\,,\ dr_2 = h_2\, du_2\,,\ dr_3 = h_3\, du_3 \tag{2.68}$$

are the basis line elements. If the position vector sweeps a curve (which is typical for example in kinematics) the *arc length element* denoted by ds is obtained by

$$(ds)^2 = d\boldsymbol{r} \cdot d\boldsymbol{r} = h_1^2 \, (du_1)^2 + h_2^2 \, (du_2)^2 + h_3^2 \, (du_3)^2 \; . \tag{2.69}$$

Note that the above formulas are completely general. For instance in Cartesian coordinates where $h_i = 1$ they are reduced to familiar forms

$$d\boldsymbol{r} = dx_1 \boldsymbol{e}_1 + dx_2 \boldsymbol{e}_2 + dx_3 \boldsymbol{e}_3 \quad \text{and} \quad (ds)^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 \; .$$

Area elements

An area element is a vector described by its magnitude and direction

$$d\boldsymbol{S} = dS \, \boldsymbol{n} \, , \tag{2.70}$$

whose decomposition is obtained by

$$d\boldsymbol{S} = (d\boldsymbol{S} \cdot \boldsymbol{e}_1) \, \boldsymbol{e}_1 + (d\boldsymbol{S} \cdot \boldsymbol{e}_2) \, \boldsymbol{e}_2 + (d\boldsymbol{S} \cdot \boldsymbol{e}_3) \, \boldsymbol{e}_3 \tag{2.71}$$

where $d\boldsymbol{S} \cdot \boldsymbol{e}_i$ is in fact the projection of $d\boldsymbol{S}$ on the coordinate surface normal to $\boldsymbol{e}_i$.

On the other hand, having the coordinate line elements (2.68), the area elements on each coordinate surface are obtained by

$$\begin{aligned} dS_1 \boldsymbol{e}_1 &= dr_2 \boldsymbol{e}_2 \times dr_3 \boldsymbol{e}_3 = h_2 \, h_3 \, du_2 \, du_3 \boldsymbol{e}_1 \\ dS_2 \boldsymbol{e}_2 &= dr_3 \boldsymbol{e}_3 \times dr_1 \boldsymbol{e}_1 = h_3 \, h_1 \, du_3 \, du_1 \boldsymbol{e}_2 \\ dS_3 \boldsymbol{e}_3 &= dr_1 \boldsymbol{e}_1 \times dr_2 \boldsymbol{e}_2 = h_1 \, h_2 \, du_1 \, du_2 \boldsymbol{e}_3 \end{aligned} \tag{2.72}$$

which are called basis area elements. Note that an area element is a parallelogram with its side being line elements, and its area is obtained by cross product of the line element vectors. Comparing the two above equation gives

$$d\boldsymbol{S} = (h_2 \, h_3 \, du_2 \, du_3) \, \boldsymbol{e}_1 + (h_3 \, h_1 \, du_3 \, du_1) \, \boldsymbol{e}_2 + (h_1 \, h_2 \, du_1 \, du_2) \, \boldsymbol{e}_3 \tag{2.73}$$

As the simplest case, in Cartesian coordinates the latter formula gives the familiar form

$$d\boldsymbol{S} = dx_2 \, dx_3 \, \boldsymbol{e}_1 + dx_3 \, dx_1 \, \boldsymbol{e}_2 + dx_1 \, dx_2 \, \boldsymbol{e}_3 \, .$$

Volume element

A volume element dV is a parallelepiped built by the three coordinate line elements $dr_1 \boldsymbol{e}_1$, $dr_2 \boldsymbol{e}_2$ and $dr_3 \boldsymbol{e}_3$. Since the volume of a parallelepiped is obtained by triple product of its side vectors as in equation (1.77), we have

$$dV = (dr_1 \boldsymbol{e}_1) \cdot (dr_2 \boldsymbol{e}_2) \times (dr_3 \boldsymbol{e}_3) = h_1 \, h_2 \, h_3 \, du_1 \, du_2 \, du_3 \, . \tag{2.74}$$

2.14 Differential transformation

For two given orthogonal curvilinear coordinate (u_1, u_2, u_3) and (q_1, q_2, q_3), we are interested in transformation of differentials between them. Here we adopt the index notation. There are two types of expressions in the aforementioned formulas to be transformed. One is the differentials du_i and the other is the factors of the form $\partial \boldsymbol{r}/\partial u_i$ or their norms h_i. We consider the transformation of each separately. Starting from transformations of the form $\boldsymbol{u}(\boldsymbol{q})$, applying the chain rule we have

$$du_i = \frac{\partial u_i}{\partial q_j} dq_j \qquad \text{and} \qquad \frac{\partial \boldsymbol{r}}{\partial u_i} = \frac{\partial \boldsymbol{r}}{\partial q_j}\frac{\partial q_j}{\partial u_i} \tag{2.75}$$

and because of orthogonality of coordinates

$$\left\| \frac{\partial \boldsymbol{r}}{\partial u_i} \right\| = \left[\sum_j \left(\frac{\partial \boldsymbol{r}}{\partial q_j}\frac{\partial q_j}{\partial u_i} \right)^2 \right]^{1/2} . \tag{2.76}$$

Transformation of area element form one curvilinear coordinates to another is obtained by

$$d\boldsymbol{A}_q = \left| \frac{\partial (q_1, q_2, q_3)}{\partial (u_1, u_2, u_3)} \right| \left(\frac{\partial (q_1, q_2, q_3)}{\partial (u_1, u_2, u_3)} \right)^{-T} \cdot d\boldsymbol{A}_u \tag{2.77}$$

And volume element transforms according

$$dV_q = \left| \frac{\partial (q_1, q_2, q_3)}{\partial (u_1, u_2, u_3)} \right| dV_u \tag{2.78}$$

2.15 Differential operators in curvilinear coordinates

Here we briefly address the general formulations of gradient, divergence, curl and Laplacian in orthogonal curvilinear coordinates. Formulations are given in a concrete form for a scalar field ϕ and a vector field $\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3$ however the reader should be able to use them for tensor fields of any order.

$$\nabla\phi = \frac{1}{h_1}\frac{\partial \phi}{\partial u_1} + \frac{1}{h_2}\frac{\partial \phi}{\partial u_2} + \frac{1}{h_3}\frac{\partial \phi}{\partial u_3} \tag{2.79}$$

$$\nabla \cdot \boldsymbol{a} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}(h_2 h_3 a_1) + \frac{\partial}{\partial u_2}(h_3 h_1 a_2) + \frac{\partial}{\partial u_3}(h_1 h_2 a_3) \right] \tag{2.80}$$

$$\nabla \times \boldsymbol{a} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} h_1\boldsymbol{e}_1 & h_2\boldsymbol{e}_2 & h_3\boldsymbol{e}_3 \\ \partial/\partial u_1 & \partial/\partial u_2 & \partial/\partial u_3 \\ h_1 a_1 & h_2 a_2 & h_3 a_3 \end{vmatrix} \tag{2.81}$$

$$\Delta\phi = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial \phi}{\partial u_1} \right) + \frac{\partial}{\partial u_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial \phi}{\partial u_2} \right) + \frac{\partial}{\partial u_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial \phi}{\partial u_3} \right) \right] . \tag{2.82}$$

Let us apply these formulas for the two most commonly used curvilinear coordinate systems, namely, cylindrical and spherical coordinates.

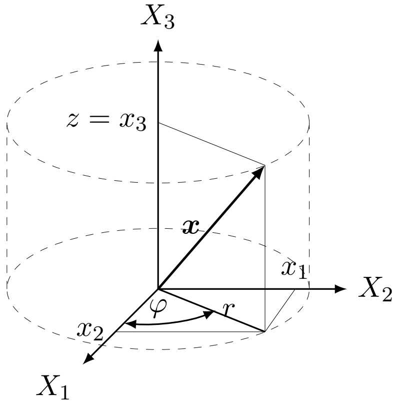

Fig. 2.7: Cylindrical coordinates.

Cylindrical coordinates

In cylindrical coordinates a point is determined by (r, φ, z) (Fig. 2.7). Transformation to rectangular coordinates are given as

$$x_1 = r\cos\varphi\,,\ x_2 = r\sin\varphi\,,\ x_3 = z \tag{2.83}$$

$$r = \sqrt{x_1^2 + x_2^2}\,,\ \varphi = \arctan(x_2/x_1)\ ,\ z = x_3 \tag{2.84}$$

Scale factors are

$$h_r = 1 \qquad h_\varphi = r \qquad h_z = 1\,, \tag{2.85}$$

and differential elements

$$d\boldsymbol{r} = dr\,\boldsymbol{e}_r + r\,d\varphi\,\boldsymbol{e}_\varphi + dz\,\boldsymbol{e}_z \tag{2.86}$$

$$d\boldsymbol{S} = r\,d\varphi\,dz\,\boldsymbol{e}_r + dr\,dz\,\boldsymbol{e}_\varphi + r\,dr\,d\varphi\,\boldsymbol{e}_z \tag{2.87}$$

$$dV = r\,dr\,d\varphi\,dz\,. \tag{2.88}$$

Finally the differential operators can be listed as

$$\nabla \equiv \boldsymbol{e}_r\frac{\partial}{\partial r} + \boldsymbol{e}_\varphi\frac{1}{r}\frac{\partial}{\partial\varphi} + \boldsymbol{e}_z\frac{\partial}{\partial z} \tag{2.89}$$

$$\nabla\cdot\boldsymbol{a} = \frac{1}{r}\frac{\partial}{\partial r}(ra_r) + \frac{1}{r}\frac{\partial a_\varphi}{\partial\varphi} + \frac{\partial a_z}{\partial z} \tag{2.90}$$

$$\Delta\phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\varphi^2} + \frac{\partial^2\phi}{\partial z^2} = \frac{\partial^2\phi}{\partial r^2} + \frac{1}{r}\frac{\partial\phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\varphi^2} + \frac{\partial^2\phi}{\partial z^2} \tag{2.91}$$

$$\nabla\times\boldsymbol{a} = \left[\frac{1}{r}\frac{\partial a_z}{\partial\varphi} - \frac{\partial a_\varphi}{\partial z}\right]\boldsymbol{e}_r + \left[\frac{\partial a_r}{\partial z} - \frac{\partial a_z}{\partial r}\right]\boldsymbol{e}_\varphi + \frac{1}{r}\left[\frac{\partial}{\partial r}(ra_\varphi) - \frac{\partial a_r}{\partial\varphi}\right]\boldsymbol{e}_z \tag{2.92}$$

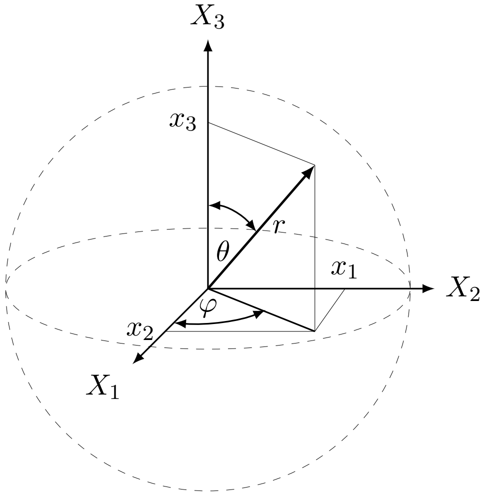

Fig. 2.8: Spherical coord.

Spherical coordinates

In spherical coordinates a point is determined by (r, θ, ϕ) (Fig. 2.8). Transformation to rectangular coordinates are given as

$$x_1 = r\sin\theta\cos\phi\,,\ x_2 = r\sin\theta\sin\phi\,,\ x_3 = r\cos\theta \tag{2.93}$$

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2}\,,\ \theta = \arccos(x_3/r)\ ,\ \phi = \arctan(x_2/x_1) \quad .$$

Scale factors are

$$h_r = 1 \qquad h_\theta = r \qquad h_\phi = r\sin\theta\,, \tag{2.94}$$

and differential elements

$$d\boldsymbol{r} = dr\,\boldsymbol{e}_r + r\,d\theta\,\boldsymbol{e}_\theta + r\sin\theta\,d\phi\,\boldsymbol{e}_\phi \tag{2.95}$$

$$d\boldsymbol{S} = r^2\sin\theta\,d\theta\,d\phi\,\boldsymbol{e}_r + r\sin\theta\,dr\,d\phi\,\boldsymbol{e}_\theta + r\,dr\,d\theta\,\boldsymbol{e}_\phi \tag{2.96}$$

$$dV = r^2\sin\theta\,dr\,d\theta\,d\phi\,. \tag{2.97}$$

Finally the differential operators can be listed as

$$\nabla \equiv \boldsymbol{e}_r\frac{\partial}{\partial r} + \boldsymbol{e}_\theta\frac{1}{r}\frac{\partial}{\partial\theta} + \boldsymbol{e}_\phi\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi} \tag{2.98}$$

$$\nabla\cdot\boldsymbol{a} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 a_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta a_\theta) + \frac{1}{r\sin\theta}\frac{\partial a_\phi}{\partial\phi} \tag{2.99}$$

$$\Delta\xi = \frac{\partial^2\xi}{\partial r^2} + \frac{2}{r}\frac{\partial\xi}{\partial r} + \frac{\cos\theta}{r^2\sin\theta}\frac{\partial\xi}{\partial\theta} + \frac{1}{r^2}\frac{\partial^2\xi}{\partial\theta^2} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\xi}{\partial\phi^2} \tag{2.100}$$

$$\nabla\times\boldsymbol{a} = \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta a_\phi) - \frac{\partial a_\theta}{\partial\phi}\right]\boldsymbol{e}_r + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial a_r}{\partial\phi} - \frac{\partial}{\partial r}(r a_\phi)\right]\boldsymbol{e}_\theta \tag{2.101}$$

$$+\frac{1}{r}\left[\frac{\partial}{\partial r}(r a_\theta) - \frac{\partial a_r}{\partial\theta}\right]\boldsymbol{e}_\phi \tag{2.102}$$

Bibliography

[1] A.I. Borisenko and I.E. Tarapov. *Vector and Tensor Analysis with Applications.* Dover Pub. Inc, 1968.

[2] P.C. Chou and N.J. Pagano. *Elasticity: Tensor, Dyadic, and Engineering Approaches.* Dover Pub. Inc, 1992.

[3] R.W. Ogden. *Non-Linear Elastic Deformations.* Dover Pub. Inc, 1997.

[4] M.R. Spiegel. *Mathematical Handbook of Formulas and Tables.* McGraw-Hill (Schaum's outline), 1979.