Introduction to vector and tensor analysis

Jesper Ferkinghoff-Borg

September 6, 2007

Contents

1 Physical space **3**
- 1.1 Coordinate systems . . . 3
- 1.2 Distances . . . 3
- 1.3 Symmetries . . . 4

2 Scalars and vectors **5**
- 2.1 Definitions . . . 5
- 2.2 Basic vector algebra . . . 5
  - 2.2.1 Scalar product . . . 6
  - 2.2.2 Cross product . . . 7
- 2.3 Coordinate systems and bases . . . 7
  - 2.3.1 Components and notation . . . 9
  - 2.3.2 Triplet algebra . . . 10
- 2.4 Orthonormal bases . . . 11
  - 2.4.1 Scalar product . . . 11
  - 2.4.2 Cross product . . . 12
- 2.5 Ordinary derivatives and integrals of vectors . . . 13
  - 2.5.1 Derivatives . . . 14
  - 2.5.2 Integrals . . . 14
- 2.6 Fields . . . 15
  - 2.6.1 Definition . . . 15
  - 2.6.2 Partial derivatives . . . 15
  - 2.6.3 Differentials . . . 16
  - 2.6.4 Gradient, divergence and curl . . . 17
  - 2.6.5 Line, surface and volume integrals . . . 20
  - 2.6.6 Integral theorems . . . 24
- 2.7 Curvilinear coordinates . . . 26
  - 2.7.1 Cylindrical coordinates . . . 27
  - 2.7.2 Spherical coordinates . . . 28

3 Tensors **30**
- 3.1 Definition . . . 30
- 3.2 Outer product . . . 31
- 3.3 Basic tensor algebra . . . 31

3.3.1 Transposed tensors . . . 32
3.3.2 Contraction . . . 33
3.3.3 Special tensors . . . 33
3.4 Tensor components in orthonormal bases . . . 34
3.4.1 Matrix algebra . . . 35
3.4.2 Two-point components . . . 38
3.5 Tensor fields . . . 38
3.5.1 Gradient, divergence and curl . . . 39
3.5.2 Integral theorems . . . 40

4 Tensor calculus **42**
4.1 Tensor notation and Einsteins summation rule . . . 43
4.2 Orthonormal basis transformation . . . 44
4.2.1 Cartesian coordinate transformation . . . 45
4.2.2 The orthogonal group . . . 45
4.2.3 Algebraic invariance . . . 46
4.2.4 Active and passive transformation . . . 47
4.2.5 Summary on scalars, vectors and tensors . . . 48
4.3 Tensors of any rank . . . 49
4.3.1 Introduction . . . 49
4.3.2 Definition . . . 50
4.3.3 Basic algebra . . . 51
4.3.4 Differentiation . . . 53
4.4 Reflection and pseudotensors . . . 54
4.4.1 Levi-Civita symbol . . . 55
4.4.2 Manipulation of the Levi-Civita symbol . . . 56
4.5 Tensor fields of any rank . . . 57

5 Invariance and symmetries **60**

Chapter 1

Physical space

1.1 Coordinate systems

In order to consider mechanical -or any other physical- phenomena it is necessary to choose a *frame of reference*, which is a set of rules for ascribing numbers to physical objects. The physical space has three dimensions, which means that it takes exactly three numbers -say x_1, x_2, x_3- to locate a point, P, in space. These numbers are called *coordinates* of a point, and the reference frame for the coordinates is called the coordinate system $\mathcal{C}$. The coordinates of P can conveniently be collected into a a triplet, $(\underline{x})_{\mathcal{C}}(P) = (x_1(P), x_2(P), x_3(P))_{\mathcal{C}}$, called the *position* of the point. Here, we use an underline simply as a short hand notation for a triplet of quantities. The same point will have a different position $(\underline{x}')_{\mathcal{C}'}(P) = (x'_1(P), x'_2(P), x'_3(P))_{\mathcal{C}'}$ in a different coordinate system $\mathcal{C}'$. However, since difference coordinate systems are just different ways of representing the same physical space in terms of real numbers, the coordinates of P in C' must be a unique function of its coordinates in C. This functional relation can be expressed as

$$\underline{x}' = \underline{x}'(\underline{x}) \tag{1.1}$$

which is a shorthand notation for

$$\begin{array}{ll} x'_1 & = x'_1(x_1, x_2, x_3) \\ x'_2 & = x'_2(x_1, x_2, x_3) \\ x'_3 & = x'_3(x_1, x_2, x_3) \end{array} \tag{1.2}$$

This functional relation is called a *coordinate transformation*. The inverse transformation must also exist, because both systems are assumed to cover the same space.

1.2 Distances

.

We will take it for given that one can ascribe a unique value (ie. independent of the chosen coordinate system), $\mathcal{D}(O,P)$, for the distance between two arbitraily chosen points O and P. Also we assume (or we may take it as an observational fact) that when $\mathcal{D}$ becomes sufficiently small, a coordinate system exists by which $\mathcal{D}$ can be calculated according to Pythagoras law:

$$\mathcal{D}(O,P) = \sqrt{\sum_{i=1}^{d}(x_i(P) - x_i(O))^2}, \quad \text{Cartesian coordinate system,} \tag{1.3}$$

where $d = 3$ is the dimension.

We will call such a coordinate system a *Cartesian* or a *Rectangular* coordinate system. Without loss of generality we can choose one of the points as origen, -say O-, so $\underline{x}(O) = (0,0,0)$. The neighborhood, $\Omega(O)$, around O for which distances between any two points, $P_1, P_2 \in \Omega(O)$ can be calculated according to Eq. (1.3) for the same fixed coordinate system is coined a (local) Euklidean space. In an Euklidean space, we also write $|P_1P_2|$ for $\mathcal{D}(P_1,P_2)$, representing the length of the line segment connecting P_1 and P_2. Trivially, $O \in \Omega(O)$.

Classical physics takes place in a 3-dimensional globally Euclidean space $\Omega(O) = \mathbb{R}^3$. In general relativity space are intrinsically curved and the assumption of an Euklidean space can only be applied locally. The principle of curved space is easier to envisage for 2d-surfaces. For instance all points on the surface of earth can be represented by two numbers -say x_1, x_2- representing the longitude and latitude. However, the law of Pythagoras (with $d = 2$) can only be applied for small rectangular triangles[1] on the surface, ie. locally. For larger rectangular triangles the sum of the angles will be larger than 180^0 and Pythagoras' law will not be correct. All geometric analysis, however, rely on the assumption that at sufficently small scales the space will appear flat. In differential geometry one only requires flatness in a differential sence.

1.3 Symmetries

For classical mechanical phenomena it is found that a frame of reference can always be chosen in which space is homogeneous and isotropic and time is homogeneous. It implies that nature has no sence of an absolute origin in time and space and no sence of an absolute direction. Mechanical or geometrical laws should therefore be invariant to coordinate transformations involving rotations and translations. Motivated by this fact we are inclined to develop a mathematical formalism for operating with geometrical objects without referring to a coordinate system or, equivalently, a formalism *which will take the same form (ie. is covariant) for all coordinates*. This fundamental principle is called the principle of covariance. The mathematics of scalar, vector and tensor algebra is precisely such a formalism.

[1] Small would mean that the length of line segments are much smaller than the radius of earth

Chapter 2

Scalars and vectors

2.1 Definitions

A vector is a quantity having both magnitude and a direction in space, such as displacement, velocity, force and acceleration.

Graphically a vector is represented by an arrow OP from a point O to a point P, defining the direction and the magnitude of the vector being indicated by the length of the arrow. Here, O is called the initial point and P is called the terminal point. Analytically, the vector is represented by either $\vec{OP}$ or $\mathbf{OP}$ and the magnitude by $|\vec{OP}|$ or $|\mathbf{OP}|$. We shall use the bold face notation in these notes. In this chapter will assume that all points P belong to an Euklidean space, $P \in \Omega(O)$, meaning that lengths of line segments can be calculated according to Pythagoras.

A scalar is a quantity having magnitude but no direction, e.g. mass, length, time, temperature and any real number.

We indicate scalars by letters of ordinary types. For example, a vector $\mathbf{a}$ will have length $a = |\mathbf{a}|$. Operations with scalars follow the same rules as elementary algebra; so multiplication, addition and substraction (provided the scalars have same units) follow the usual algebraic rules.

2.2 Basic vector algebra

The operations defined for real numbers are, with suitable definitions, capable of extension to an algebra of vectors. The following definitions are fundamental and define the basic algebraic rules of vectors:

1. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are equal if they have the same magnitude and direction regardless of the position of their initial point.

2. A vector having direction opposite of a vector $\mathbf{a}$ but having the same magnitude is denoted $-\mathbf{a}$.

3. The *sum* of *resultant* of vectors $\mathbf{a}$ and $\mathbf{b}$ is a vector $\mathbf{c}$ formed by placing the initial point of $\mathbf{b}$ on the terminal point of $\mathbf{a}$ and then joining the initial point of $\mathbf{a}$ to the terminal point of $\mathbf{b}$. The sum is written $\mathbf{c} = \mathbf{a} + \mathbf{b}$.

4. The *difference* between two vectors, $\mathbf{a}$ and $\mathbf{b}$, represented by $\mathbf{a} - \mathbf{b}$ is defined as the sum $\mathbf{a} + (-\mathbf{b})$.

5. The *product* of a vector $\mathbf{a}$ with a scalar m is a vector $m\mathbf{a}$ with magnitude $|m|$ times the magnitude of $\mathbf{a}$, ie $|m\mathbf{a}| = |m||\mathbf{a}|$, and with direction the same as or opposite to that of $\mathbf{a}$, according as m is positive or negative.

We stress that these definitions for vector addition, substraction and scalar multiplications are defined *geometrically*, ie. they have no reference to coordinates. It then becomes a pure geometric exercise to prove the following laws:

$$\begin{array}{rll} \mathbf{a}+\mathbf{b} & = \mathbf{b}+\mathbf{a} & \text{Commutative law for addition} \\ \mathbf{a}+(\mathbf{b}+\mathbf{c}) & = (\mathbf{a}+\mathbf{b})+\mathbf{c} & \text{Associate law for addition} \\ m(n\mathbf{a}) & = (mn)\mathbf{a} & \text{Associate law for multiplication} \\ (m+n)\mathbf{a} & = m\mathbf{a}+n\mathbf{a} & \text{Distributive law} \\ m(\mathbf{a}+\mathbf{b}) & = m\mathbf{a}+n\mathbf{b} & \text{Distributive law} \\ \mathbf{a}m & =_{\text{def}} \mathbf{a}m & \text{Commutative law for multiplication} \end{array} \tag{2.1}$$

Here, the last formula should just be read as a definition of a vector times a scalar. Note that in all cases only multiplication of a vector by one or more scalars are defined. One can define different types of bilinear vector products. The three basic types are called *scalar product* (or inner product), *cross product* and *outer product* (or tensor product). We shall define each in turn. The definition of the outer product is postponed to chapter 3.

Following definition will become useful:

A *unit vector* is a vector having unit magnitude. If $\mathbf{a}$ is not a null vector then $\mathbf{a}/|\mathbf{a}|$ is a unit vector having the same direction as $\mathbf{a}$.

2.2.1 Scalar product

The scalar product between two vectors, $\mathbf{a}$ and $\mathbf{b}$ is defined by

$$\mathbf{a} \cdot \mathbf{b} = ab\cos(\theta), \quad 0 \leq \theta \leq \pi \tag{2.2}$$

where $a = |\mathbf{a}|$, $b = |\mathbf{b}|$ and θ is the angle between the two vectors. Note that $\mathbf{a} \cdot \mathbf{b}$ is a scalar. Following rules apply:

$$\begin{array}{rll} \mathbf{a}\cdot\mathbf{b} & = \mathbf{b}\cdot\mathbf{a} & \text{Commutative law for scalar product} \\ (\mathbf{a}+\mathbf{b})\cdot\mathbf{c} & = \mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c} & \text{Distributive law in 1. argument} \\ \mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) & = \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c} & \text{Distributive law for 2. argment} \\ m(\mathbf{a}\cdot\mathbf{b}) & = (m\mathbf{a})\cdot\mathbf{b} = \mathbf{a}\cdot(m\mathbf{b}) & m \text{ is a scalar} \end{array} \tag{2.3}$$

The three last formulas make the scalar product *bilinear* in the two arguments. Note also,

$$\begin{array}{ll} 1 & \mathbf{a}\cdot\mathbf{a} = |\mathbf{a}|^{\mathbf{2}} \\ 2 & \text{If } \mathbf{a}\cdot\mathbf{b} = \mathbf{0} \text{ and } \mathbf{a} \text{ and } \mathbf{b} \text{ are not null vectors,} \\ & \text{then } \mathbf{a} \text{ and } \mathbf{b} \text{ are perpendicular.} \\ 3 & \text{The projection of a vector } \mathbf{a} \text{ on } \mathbf{b} \text{ is equal to } \mathbf{a}\cdot\mathbf{e_b}, \\ & \text{where } \mathbf{e_b} = \mathbf{b}/|\mathbf{b}| \text{ is the unit vector in direction of } \mathbf{b}. \end{array} \tag{2.4}$$

2.2.2 Cross product

The cross product, $\mathbf{a}\times\mathbf{b}$ between two vectors $\mathbf{a}$ and $\mathbf{b}$ is a vector defined by

$$\mathbf{a}\times\mathbf{b} = ab\sin(\theta)\mathbf{u}, \quad 0\le\theta\le\pi, \tag{2.5}$$

where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$ and $\mathbf{u}$ is a unit vector in the direction perpendicular to the plane of $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a},\mathbf{b}$ and $\mathbf{u}$ form a right-handed system[1].

The following laws are valid:

$$\begin{array}{rll} \mathbf{a}\times\mathbf{b} & = -\mathbf{b}\times\mathbf{a} & \text{Cross product is } \textit{not} \text{ commutative.} \\ (\mathbf{a}+\mathbf{b})\times\mathbf{c} & = \mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c} & \text{Distributive law for 1. argument} \\ \mathbf{a}\times(\mathbf{b}+\mathbf{c}) & = \mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c} & \text{Distributive law for 2. argument} \\ m(\mathbf{a}\times\mathbf{b}) & = (m\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(m\mathbf{b}) & m \text{ is a scalar} \end{array} \tag{2.6}$$

The last three formulas make the cross product bilinear.

Note that:

1. The absolute value of the cross product $|\mathbf{a}\times\mathbf{b}|$ has a particular geometric meaning. It equals the area of the parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$.

2. The absolute value of the triple product $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ has a particular geometric meaning. It equals the volume of the parallelepiped spanned by $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$

2.3 Coordinate systems and bases

We emphasize again that all definitions and laws of vector algebra, as introduced above, are invariant to the choise of coordinate system[2]. Once we introduce a way of ascribing positions to points by the choise of a coordinate system, $\mathcal{C}$, we obtain a way of *representing vectors in terms of triplets*.

[1] It is seen that the definition of the cross-product explicitly depends on an arbitrary choise of *handedness*. A vector whose direction depends on the choise of handedness is called an *axial vector* or *pseudovector* as opposed to ordinary or *polar vectors*, whose directions are independent of the choise of handedness.

[2] With the one important exception that we have assumed the absolute notion of handedness in the definition of the cross product.

By assumption, we can choose this coordinate system to be rectangular (R), $\mathcal{C}_R$, cf. chapter 1. By selecting four points O, P_1, P_2 and P_3 having the coordinates

$$\underline{x}_{\mathcal{C}_R}(O) = (0,0,0),\ \ \underline{x}_{\mathcal{C}_R}(P_1) = (1,0,0),\ \ \underline{x}_{\mathcal{C}_R}(P_2) = (0,1,0),\ \ \underline{x}_{\mathcal{C}_R}(P_3) = (0,0,1) \tag{2.7}$$

we can construct three vectors

$$\mathbf{e_1} = \mathbf{OP_1},\ \ \mathbf{e_2} = \mathbf{OP_2},\ \ \mathbf{e_3} = \mathbf{OP_3}. \tag{2.8}$$

Furthermore, from the choise of an origin O, a *position vector*, $\mathbf{r} = \mathbf{OP}$, can be associated to any point P by the formula

$$\mathbf{r}(P) = \mathbf{r}(\underline{x}_{\mathcal{C}_R}(P)) = \sum_i x_i(P)\mathbf{e}_i \qquad \mathcal{C}_R \ \text{ rectangular.} \tag{2.9}$$

The position vector is an *improper* vector, meaning that it explicitly depends on the position of its initial point O. It will therefore change upon a transformation to a coordinate system with a different origin. However, we shall see that all physical laws only involve positional displacements,

$$\Delta\mathbf{r} = \mathbf{r}(P_2) - \mathbf{r}(P_1) = \mathbf{OP_2} - \mathbf{OP_1} = \mathbf{P_1P_2},$$

which is a vector invariant to the choise of the coordinate system.

Any point will have one and only one associated position vector by Eq. (2.9). Thus, the three vectors $\mathcal{R} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ form a *basis* for the vector space[3]. In any coordinate system, $\mathcal{C}$, a well-defined procedure exists for constructing the associated basis, $\mathcal{G}$, of the vector space[4]. However, there are two definining characteristica for a rectangular system

1. The basis is *spatially constant*, ie. the basis vectors are not themselves a function of position.

2. The basis is *orthonormal*, meaning that the basis vectors are mutually orthogonal, $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ for $i \neq j$, and unitary $|\mathbf{e}_i| = 1$.

The basis $\mathcal{G} = \{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ associated with a general coordinate system, $\mathcal{C}$, may satisfy none of the above. There will always be a one to one relation between the coordinates and the position vector in geometrical problems of classical physics, but it is only for bases satisfying 1, that this relation will be on the form of Eq. (2.9) with the replacement $\mathbf{e}_i \to \mathbf{g}_i$. In thes cases, $\mathcal{G}$ and the choise of origin, O, fully specifies the coordinate system, $\mathcal{C} = (O, \mathcal{G})$.

[3] Recall that the definition of a basis is a set of vectors, $\mathcal{G}$, that span the vector space and are linear independent. A set of vectors will be a basis if and only if the triple product $\mathbf{g}_1 \cdot (\mathbf{g}_2 \times \mathbf{g}_3) \neq 0$.

[4] The simple procedure of Eq. (2.7) and (2.8) for obtaining the basis vectors is not generally applicable. We shall illustrate the general method in section (2.7).

In the following we shall retain the notation $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for the particular case of an orthonormal basis[5]. If $\mathbf{e}_i$ is non-constant we explicitly write $\mathbf{e}_i(\underline{x})$. On the other hand, if $\mathcal{E}$ is constant the coordinate system is rectangular and we use the letter $\mathcal{R}$ for the basis.

2.3.1 Components and notation

From the definition of any basis, $\mathcal{G}$, spatially dependent or not, an abitrary vector $\mathbf{a}$ will have a unique triplet of quantities $\underline{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ with respect to $\mathcal{G}$ such that

$$\mathbf{a} = \sum_i a_i \mathbf{g}_i. \tag{2.10}$$

These quantities are called the *components* of $\mathbf{a}$ with respect to $\mathcal{G}$. Mathematically speaking, we obtain a one-to-one map $\phi_{\mathcal{G}} : V \to \mathbb{R}^3$ between the vector space, V, and the triplet of scalars $\mathbb{R}^3$, with the choise of a basis. It is useful to have a short-hand notation for this map (ϕ) and its inverse (ϕ^{-1}), ie. for the resolvent of a vector into its components and for the vector obtained by expanding a set of components along the basis vectors. We define the the following two bracket functions for the two operations respectively:

$$\begin{array}{lll} [\mathbf{a}]_{\mathcal{G}} & =_{\text{def.}} \phi_{\mathcal{G}}(\mathbf{a}) & = \underline{a} \qquad\qquad\quad \in \mathbb{R}^3 \\ (\underline{a})_{\mathcal{G}} & =_{\text{def.}} \phi_{\mathcal{G}}^{-1}(\underline{a}) & = \sum_i a_i \mathbf{g}_i = \mathbf{a} \quad \in V \end{array}$$

The notation for the set of components of a vector $\mathbf{a}$, $\underline{a} = [\mathbf{a}]_{\mathcal{G}}$, also mean that we shall refer to the i''th component as $a_i = [\mathbf{a}]_{\mathcal{G},i}$. Furthermore, when it is obvious from the context (or irrelevant) which basis is implied, we will omit the basis subscript and use the notation $\underline{a} = [\mathbf{a}]$ and $a_i = [\mathbf{a}]_i$. For the collection of components, a_i, into a triplet, ordinary brackets are commonly used, so

$$(a_i) =_{\text{def.}} (a_1, a_2, a_3) = \underline{a}.$$

In order not to confuse this notation with the vector obtained from a set of components

$$(\underline{a})_{\mathcal{G}} =_{\text{def.}} \sum_i a_i \mathbf{g}_i,$$

whe shall always retain the basis subscript in the latter expression.

Trivially, we have

$$\mathbf{g_1} = (1,0,0)_{\mathcal{G}}, \quad \mathbf{g_2} = (0,1,0)_{\mathcal{G}}, \quad \mathbf{g_3} = (0,0,1)_{\mathcal{G}}$$

or in the more dense, index-based notation[6]

$$[\mathbf{g_i}]_j = \delta_{ij}, \qquad \text{wrt. basis } \mathcal{G} = \{\mathbf{g_1}, \mathbf{g_2}, \mathbf{g_3}\}$$

[5] This basis need not be spatially independent. In the case of a constant orthonormal basis, ie. a Cartesian basis, the notation $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is also often used. However, we shall retain the other notation for algebraic convenience.

[6] Expressing vector identities in terms of their components is also referred to as *tensor notation*.

where δ_{ij} is the Kronecker delta

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{2.11}$$

Furthermore, if $\mathcal{G}$ is constant in space the relation between coordinates and the position vector is on the form of Eq. (2.9)

$$\mathbf{r}(\underline{x}) = (\underline{x})_{\mathcal{G}} = \sum_i x_i \mathbf{g}_i \quad \Leftrightarrow \quad [\mathbf{r}]_{\mathcal{G},i} = x_i \quad \mathcal{G} \text{ const.} \tag{2.12}$$

With the risk of being pedantic let us stress the difference between a vector and its components. For two different bases $\mathcal{G} \neq \mathcal{G}'$, a vector $\mathbf{a}$ will be represented by the components $\underline{a} \in \mathbb{R}^3$ and $\underline{a}' \in \mathbb{R}^3$ respectively, so that

$$\mathbf{a} = (\underline{a})_{\mathcal{G}} = (\underline{a}')_{\mathcal{G}'} \quad but \quad \underline{a} \neq \underline{a}'.$$

Though representing the same vector, the two sets of components $\underline{a}$ and $\underline{a}'$ will differ because they refer to different bases. The distinction between a vector and its components becomes irrelevant in the case where one basis is involved only.

2.3.2 Triplet algebra

By choosing a basis $\mathcal{G}$ for the vector space, the basic algebraic vector laws of addition, substraction and scalar multiplication, Eq. (2.1), can be expressed in terms of normal triplet-algebra for the components. For instance, since

$$\mathbf{a} + \mathbf{b} = \sum_i a_i \mathbf{g_i} + \sum_i b_i \mathbf{g_i} = \sum_i (a_i + b_i) \mathbf{g_i}$$

we have

$$\mathbf{a} + \mathbf{b} = (a_1, a_2, a_3)_{\mathcal{G}} + (b_1, b_2, b_3)_{\mathcal{G}} = (a_1 + b_1, a_2 + b_2, a_3 + b_3)_{\mathcal{G}},$$

or equivalently (for all i)

$$[\mathbf{a} + \mathbf{b}]_i = [\mathbf{a}]_i + [\mathbf{b}]_i \tag{2.13}$$

The same naturally holds for vector substraction. Also, since

$$m\mathbf{a} = m(\sum_i a_i \mathbf{g_i}) = \sum_i m a_i \mathbf{g_i},$$

for a scalar quantity m, we have

$$m\mathbf{a} = m(a_1, a_2, a_3)_{\mathcal{G}} = (ma_1, ma_2, ma_3)_{\mathcal{G}},$$

or

$$[m\mathbf{a}]_i = m[\mathbf{a}]_i. \tag{2.14}$$

The bilinear scalar and cross product are not as easy to operate with in terms of the components in an arbitrary basis $\mathcal{G}$ as in the case of vector addition, substraction and scalar multiplication. For instance

$$\begin{aligned} \mathbf{a}\cdot\mathbf{b} &= (a_1,a_2,a_3)_{\mathcal{G}}\cdot(b_1,b_2,b_3)_{\mathcal{G}} \\ &= (\textstyle\sum_i a_i\mathbf{g_i})\cdot\left(\sum_j b_j\mathbf{g_j}\right) \\ &= \textstyle\sum_{ij} a_ib_j\mathbf{g_i}\cdot\mathbf{g_j} \\ &= \textstyle\sum_{ij} a_ib_jg_{ij}, \quad g_{ij}=\mathbf{g_i}\cdot\mathbf{g_j} \end{aligned} \tag{2.15}$$

The quantities g_{ij}, are called the *metric coefficients*. If $\mathcal{G}$ is spatially dependent then $g_{ij}=g_{ij}(\underline{x})$ will be functions of the coordinates. In fact,the functional form of the metric coefficients fully specifies the type of geometry involved (Euklidean or not) and they are the starting point for extending vector calculus to arbitrary geometries. Here it suffices to say that a Cartesian coordinate system uniquely implies that the metric coefficients are constant and satisfy $g_{ij}=\delta_{ij}$.

2.4 Orthonormal bases

An orthonormal basis, $\mathcal{E}$, satisfies by definition

$$\mathbf{e_i}\cdot\mathbf{e_j}=\delta_{ij}. \tag{2.16}$$

This relation is particular handy for obtaining the components of a vector $\mathbf{a}$. Indeed, we have

$$\mathbf{a}=\textstyle\sum_i a_i\mathbf{e_i} \quad\Rightarrow\quad a_i=\mathbf{a}\cdot\mathbf{e_i}, \quad \mathcal{E} \text{ ortonormal.} \tag{2.17}$$

So we obtain the i'th component of a vector $\mathbf{a}$ by taking the scalar product between $\mathbf{a}$ and the i'th basis-vector.

2.4.1 Scalar product

In an ortonormal basis the scalar product, Eq. (2.15), is a simple expression in terms of the components

$$\mathbf{a}\cdot\mathbf{b}=(\sum_i a_i\mathbf{e_i})\cdot(\sum_j b_j\mathbf{e_j})=\sum_i a_ib_i. \tag{2.18}$$

Furthermore, if the orthonormal basis is spatially independent, the displacement vector, $\mathbf{PQ}$, between two points P and Q with coordinates $\underline{p}=(p_1,p_2,p_3)$ and $\underline{q}=(q_1,q_2,q_3)$ respectively, will be

$$\mathbf{PQ}=\mathbf{r}(Q)-\mathbf{r}(P)=\sum_i(q_i-p_i)\mathbf{e}_i$$

The distance between P and Q is given by the length of $|\mathbf{PQ}|$ and

$$|\mathbf{PQ}|^2=\sum_i(q_i-p_i)^2 \quad\Rightarrow\quad |\mathbf{PQ}|=\sqrt{\sum_i(q_i-p_i)^2}. \tag{2.19}$$

In other words, we recover the law of Pythagoras, cf. Eq. (1.3), which is reassuring since a constant orthonormal basis is equivalent to the choise of a rectangular coordinate system.

2.4.2 Cross product

Let $\mathcal{E}$ be an ortonormal right-handed basis, so $\mathbf{e_3} = \mathbf{e_1} \times \mathbf{e_2}$. Using the laws of cross-product, Eq. (2.6), one can show

$$\mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \mathbf{e_1} & \mathbf{e_2} & \mathbf{e_3} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)_{\mathcal{E}} \quad (2.20)$$

where det() denotes the determinant of the array considered as a matrix. An alternative expression defined only in terms of the components, is

$$[\mathbf{a} \times \mathbf{b}]_i = \sum_{jk} \epsilon_{ijk} a_j b_k \quad (2.21)$$

The symbol ϵ_{ijk} with the three indices is the *Levi-Civita symbol*. It consists of $3 \times 3 \times 3 = 27$ real numbers given by

$$\begin{aligned} &\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1 \\ &\epsilon_{321} = \epsilon_{132} = \epsilon_{213} = -1 \\ &\epsilon_{ijk} = 0 \quad \text{otherwise} \end{aligned} \quad (2.22)$$

In words, ϵ_{ijk} is antisymmetric in all indices. From this it follows that the ortonormal basis vectors satisfy

$$\mathbf{e_i} \times \mathbf{e_j} = \sum_k \epsilon_{ijk} \mathbf{e_k}, \quad (2.23)$$

i.e. $\mathbf{e_3} = \mathbf{e_1} \times \mathbf{e_2}$, $\mathbf{e_1} = \mathbf{e_2} \times \mathbf{e_3}$, and so on.

Applying eq. (2.3) and (2.6) one can show that the triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ in an ortonormal right-handed basis is given by

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \sum_{ijk} \epsilon_{ijk} a_i b_j c_k = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \quad (2.24)$$

The problem of handedness

Let us emphasize that Eq. (2.20,2.21,2.23,2.24) would have been the same had we defined the cross-product by a left-hand rule and expressed the components in a left-handed coordinate system. Indeed, the definition of the Levi-Civita symbol, Eq. (2.22), does not itself depend on the handedness of the coordinate system. Since it is in fact not possible to give an absolute prescribtion of how to construct a right-handed (or left-handed) coordinate system it is more common

in modern vector analysis to define the cross product through the determinant, Eq. (2.20) og equivalently through the Levi-Civita symbol, Eq. (2.21). In replacing the geometric definition, Eq. (2.5) which pressumes an absolute notion of handedness, with an algebraic one the direction of the cross-product will simply be determined by what-ever convention of handedness that has been adopted with the ortonormal coordinate system in the first place.

Physically, the arbitrariness in the choise of handedness implies that the orientation of the cross product of two ordinary vectors is not a physical objective quantity, in contrast to vectors representing displacements, velocities, forces etc. One therefore distinguishes between proper or *polar* vectors whose direction are independent on the choise of handedness and *axial* or *pseudovectors* whose directions depend on the choise of handedness. The distinction between the two types of vectors becomes important when one considers transformations between left- and right-handed coordinate systems.

Dimensionality

For all vector operations the only explicit reference to dimension of the physical space appears in the cross product. This can be seen from the definition of the ϵ-symbol, which explicitly has three indices running from 1 to 3. All other vector operations are valid in arbitrary dimensions.

The generalization of the cross product to any dimension d is obtained by constructing a Levi-Civita symbol having d indices each running from 1 to d and being totally antisymmetric.

For instance for $d = 2$:

$$\begin{aligned} \epsilon_{12} &= 1 \\ \epsilon_{21} &= -1 \\ \epsilon_{11} &= \epsilon_{22} = 0 \end{aligned} \tag{2.25}$$

Here, we get the "cross-product", $\hat{\mathbf{a}}$, of a vector $\mathbf{a} = (a_1, a_2)_{\mathcal{E}}$ by

$$[\hat{\mathbf{a}}]_j = \sum_i \epsilon_{ji} a_i$$

or in doublet notation

$$\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2)_{\mathcal{E}} = (a_2, -a_1)_{\mathcal{E}},$$

which is the familiar expression in $d = 2$ for obtaining a vector orthogonal to a given vector. Generally, in $d > 1$ dimensions the cross product will take $d - 1$ vectors as arguments.

2.5 Ordinary derivatives and integrals of vectors

Let $\mathbf{c}(t)$ be a vector depending on a single scalar variable t. In this section we will consider how to define derivatives and integrals of $\mathbf{c}$ with respect to t.

2.5.1 Derivatives

The definition for ordinary derivatives of functions can directly be extended to vector functions of scalars:

$$\frac{d\mathbf{c}(t)}{du} = \lim_{\Delta u \to 0} \frac{\mathbf{c}(u+\Delta u) - \mathbf{c}(u)}{\Delta u}. \tag{2.26}$$

Note that the definition involves substracting two vectors, $\Delta\mathbf{c} = \mathbf{c}(u+\Delta u) - \mathbf{c}(u)$ which is a vector. Therefore, $\frac{d\mathbf{r}(t)}{du}$ will also be a vector provided the limit, Eq. (2.26), exists. Eq. (2.26) is defined independently of coordinate systems. In an orthonormal basis independent of u:

$$\frac{d}{du}\mathbf{a} = \sum_i \frac{da_i}{du}\mathbf{e}_i, \tag{2.27}$$

whence the derivative of a vector is reduced to ordinary derivatives of the scalar functions a_i.

If u represents the time $u = t$ and $\mathbf{c}(t)$ is the position vector of a particle $\mathbf{r}(t)$ with respect to some coordinate system then $\frac{d\mathbf{r}(t)}{dt}$ will represent the *velocity* $\mathbf{v}(t)$. The velocity will be a tangentiel vector to the *space curve* mapped by $\mathbf{r}(t)$. Even though the position vector itself is not a proper vector due to its dependence on the choise of origin of the coordinate system, cf. section 2.3, the velocity *is* a proper vector since it is defined through the positional displacement, $\Delta\mathbf{r}$, which is a proper vector. Consequently, the *acceleration* $\mathbf{a}(t) =_{\text{def.}} \frac{d\mathbf{v}}{dt}$ is also a proper vector.

Ordinary differentiation of a vector co-exists nicely with the basic vector operations, e.g.

$$\frac{d}{du}(\mathbf{a}+\mathbf{b}) = \frac{d\mathbf{a}}{du} + \frac{d\mathbf{b}}{du} \tag{2.28}$$

$$\frac{d}{du}(\mathbf{a}\cdot\mathbf{b}) = \frac{d\mathbf{a}}{du}\cdot\mathbf{b} + \mathbf{a}\cdot\frac{d\mathbf{b}}{du} \tag{2.29}$$

$$\frac{d}{du}(\mathbf{a}\times\mathbf{b}) = \frac{d\mathbf{a}}{du}\times\mathbf{b} + \mathbf{a}\times\frac{d\mathbf{b}}{du} \tag{2.30}$$

$$\frac{d}{du}(\phi\mathbf{a}) = \frac{d\phi}{du}\mathbf{a} + \phi\frac{d\mathbf{a}}{du} \tag{2.31}$$

where $\phi(u)$ is a normal scalar function. The above formulas can of course be expressed in terms of the vector components, a_i and b_i, using Eq. (2.27).

2.5.2 Integrals

The definition of ordinary integral of a vector depending on a single scalar variable, $\mathbf{c}(u)$, also follows that of ordinary calculus. With respect to a basis $\mathcal{E}$ independent of u the *indefinite integral* of $\mathbf{c}(u)$ is defined by

$$\int \mathbf{c}(u)du = \sum_i \int c_i(u)du\mathbf{e}_i,$$

where $\int c_i(u)du$ is the indefinite integral of an ordinary scalar function. If $\mathbf{s}(u)$ is a vector satisfying $\mathbf{c}(u) = \frac{d}{du}\mathbf{s}(u)$ then

$$\int \mathbf{c}(u)du = \int \frac{d}{du}\left(\mathbf{s}(u)\right)du = \mathbf{s}(u) + \mathbf{k}$$

where $\mathbf{k}$ is an *arbitrary constant vector* independent of u. The *definite integral* between two limits $u = u_0$ and $u = u_1$ can in such case be written

$$\int_{u_0}^{u_1} \mathbf{c}(u)du = \int_{u_0}^{u_1} \frac{d}{du}(\mathbf{s}(u))du = [\mathbf{s}(u) + \mathbf{k}]_{u_0}^{u_1} = \mathbf{s}(u_1) - \mathbf{s}(u_0).$$

This integral can also be defined as a limit of a sum in a manner analogous to that of elementary integral calculus.

2.6 Fields

2.6.1 Definition

In continuous systems the basic physical variables are distributed over space. A function of space is known as a *field*. Let an arbitrary coordinate system be given.

Scalar field. If to each position $\underline{x} = (x_1, x_2, x_3)$ of a region in space the corresponds a number or scalar $\phi(x_1, x_2, x_3)$, then ϕ is called a *scalar function of position* or *scalar field*. Physical examples of scalar fields are the mass or charge density distribution of an object, the temperature or the pressure distribution at a given time in a fluid.

Vector field. If to each position $\underline{x} = (x_1, x_2, x_3)$ of a region in space there corresponds a vector $\mathbf{a}(x_1, x_2, x_3)$ then $\mathbf{a}$ is called a *vector function of position* or a *vector field*. Physical examples of vector fields are the gravitational field around the earth, the velocity field of a moving fluid, electro-magnetic field of charged particle systems.

In the following we shall assume the choise of a rectangular coordinate system $\mathcal{C} = (O, \mathcal{R})$. Introducing the position vector $\mathbf{r}(\underline{x}) = \sum_i x_i \mathbf{e}_i$ the expressions $\phi(\mathbf{r})$ and $\mathbf{a}(\mathbf{r})$ is taken to have the same meaning as $\phi(\underline{x})$ and $\mathbf{a}(\underline{x})$, respectively.

2.6.2 Partial derivatives

Let $\mathbf{a}$ be a vector field $\mathbf{a}(\mathbf{r}) = \mathbf{a}(x_1, x_2, x_3)$. We can define the partial derivative $\frac{\partial \mathbf{a}}{\partial x_i}$ in the same fashion as in Eq. (2.26). Here we will use the short hand notation:

$$\begin{aligned} \partial_i &= \frac{\partial}{\partial x_i} \\ \partial_{ij}^2 &= \frac{\partial^2}{\partial x_i \partial x_j} \end{aligned}$$

In complete analogy to the usual definition of partial derivatives of a scalar function, the partial derivative of a vector field with respect to x_i $(i = 1, 2, 3)$ is

$$(\partial_i \mathbf{a})(\mathbf{r}) = \lim_{\Delta x_i \to 0} \frac{\mathbf{a}(\mathbf{r} + \Delta x_i \mathbf{e_i}) - \mathbf{a}(\mathbf{r})}{\Delta x_i} \tag{2.32}$$

Equation (2.32) expresses how the vector field changes in the spatial direction of $\mathbf{e_i}$. For the same arguments as listed in previous section, $\partial_i \mathbf{a}$ will for each i also be a vector field. Note that in general, $\partial_j \mathbf{a} \neq \partial_i \mathbf{a}$ when $j \neq i$.

Rules for partial diffentiation of vectors are similar to those used in elementary calculus for scalar functions Thus if $\mathbf{a}$ and $\mathbf{b}$ are functions of $\underline{x}$ then

$$\begin{aligned}
\partial_i(\mathbf{a}\cdot\mathbf{b}) &= \mathbf{a}\cdot(\partial_i\mathbf{b}) + (\partial_i\mathbf{a})\cdot\mathbf{b} \\
\partial_i(\mathbf{a}\times\mathbf{b}) &= (\partial_i\mathbf{a})\times\mathbf{b} + \mathbf{a}\times(\partial_i\mathbf{b}) \\
\partial^2_{ji}(\mathbf{a}\cdot\mathbf{b}) &= \partial_j\left(\partial_i(\mathbf{a}\cdot\mathbf{b})\right) = \partial_j\left((\partial_i\mathbf{a})\cdot\mathbf{b} + \mathbf{a}\cdot(\partial_i\mathbf{b})\right) \\
&= \mathbf{a}\cdot(\partial^2_{ji}\mathbf{b}) + (\partial_j\mathbf{a})\cdot(\partial_i\mathbf{b}) + (\partial_i\mathbf{a})\cdot(\partial_j\mathbf{b}) + (\partial^2_{ji}\mathbf{a})\cdot\mathbf{b}
\end{aligned}$$

We can verify these expressions by resolving the vector fields into $\mathcal{R} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Then all differentiation is reduced to ordinary differention of the scalar functions, $a_k(\underline{x})$ and $b_k(\underline{x})$. Notice that for the position vector

$$\partial_i \mathbf{r} = \mathbf{e}_i. \tag{2.33}$$

If the vector field is resolved into a non-cartesian coordinate system, $\mathcal{C}'$

$$\mathbf{a}(\underline{x}') = \sum_j a_j(\underline{x}')\mathbf{g}_j(\underline{x}') \tag{2.34}$$

then one must remember to include the spatial derivatives of the basis vectors as well

$$\frac{\partial \mathbf{a}}{\partial x'_i}(\underline{x}') = \sum_j \Big(\frac{\partial a_j}{\partial x'_i}(\underline{x}')\mathbf{g}_j(\underline{x}') + a_j(\underline{x}')\frac{\partial \mathbf{g}_j}{\partial x'_i}(\underline{x}') \Big) \tag{2.35}$$

2.6.3 Differentials

Since partial derivatives of vector field follow those used in elementary calculus for scalar functions the same will be true for vector differentials. For example, for $\mathcal{C} = (O, \mathcal{R})$

$$\text{if} \quad \mathbf{a}(\underline{x}) = \sum_i a_i(\underline{x})\mathbf{e}_i, \quad \text{then} \quad d\mathbf{a} = \sum_i da_i(\underline{x})\mathbf{e}_i \tag{2.36}$$

The change of the vector components due to a infitesimal change of coordinates dx_j is obtained by the chain rule

$$da_i(\underline{x}) = \sum_j (\partial_j a_i)(\underline{x})dx_j \tag{2.37}$$

In vector notation it reads

$$d\mathbf{a}(\underline{x}) = \sum_i \Big(\sum_j (\partial_j a_i)(\underline{x})dx_j \Big)\mathbf{e}_i \tag{2.38}$$

or simply

$$d\mathbf{a} = \sum_i \partial_i \mathbf{a} dx_i. \tag{2.39}$$

The position vector $\mathbf{r} = \sum_i x_i \mathbf{e}_i$ is a special vector field for which Eq. (2.38) implies

$$d\mathbf{r} = \sum_i dx_i \mathbf{e}_i. \tag{2.40}$$

The square, $(ds)^2$ of the infinitesimal distance moved is then given by

$$(ds)^2 = d\mathbf{r} \cdot d\mathbf{r} = \sum_i dx_i^2$$

If $\mathbf{r}(u)$ is space curve parametrized by u then

$$\left(\frac{ds}{du}\right)^2 = \frac{d\mathbf{r}}{du} \cdot \frac{d\mathbf{r}}{du},$$

Therefore, the arc length between two points on the curve $\mathbf{r}(u)$ given by $u = u_1$ and $u = u_2$ is

$$s = \int_{u_1}^{u_2} \sqrt{\frac{d\mathbf{r}}{du} \cdot \frac{d\mathbf{r}}{du}}\, du.$$

In general, we have (irrespective of the choise of basis)

$$\begin{aligned} d(\mathbf{a} \cdot \mathbf{b}) &= d\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot d\mathbf{b} \\ d(\mathbf{a} \times \mathbf{b}) &= d\mathbf{a} \times \mathbf{b} + \mathbf{a} \times d\mathbf{b} \end{aligned}$$

2.6.4 Gradient, divergence and curl

Let a rectangular coordinate system be given $\mathcal{C} = (O, \mathcal{R})$.

Nabla operator ∇

The vector differential operator *del* or *nabla* written as ∇ is defined by

$$\nabla(\cdot) = \sum_i \mathbf{e}_i \partial_i (\cdot) \tag{2.41}$$

where $\cdot$ represents a scalar or –as we shall see later– a vector field. Notice that the i'th component of the operator is given by

$$\nabla_i = \mathbf{e}_i \cdot \nabla = \partial_i.$$

This vector operator possesses properties analogous to those of ordinary vectors which is not obvious at first sight, since its operation is defined relative to the choise of the coordinate system, in contrast to the vector algebra defined hitherto. We shall postpone the demonstration of its vectorial nature until section 4.5, where we have become more familiar with tensor calculus.

The *gradient* of a scalar field $\phi(\mathbf{r})$ is defined by inserting a scalar field in the above definition:

$$grad\ \phi = \nabla\phi = \sum_i (\partial_i \phi)\mathbf{e}_i$$

The gradient of a scalar field has some interesting geometrical properties. Consider the change of ϕ in some particular direction. For an infinitesimal vector displacement, $d\mathbf{r}$, forming its scalar product with $\nabla\phi$ we have

$$(\nabla\phi)(\mathbf{r}) \cdot d\mathbf{r} = \left(\sum_i (\partial_i\phi)(\mathbf{r})\mathbf{e}_i\right) \cdot \left(\sum_j dx_j \mathbf{e}_j\right) = \sum_i (\partial_i\phi)(\mathbf{r}) dx_i = d\phi,$$

which is the infinitesimal change in ϕ in going from from position $\mathbf{r}$ to $\mathbf{r} + d\mathbf{r}$. In particular, if $\mathbf{r}$ depends on some parameter t such that $\mathbf{r}(t)$ defines a space curve then the total derivative of ϕ with respect to that curve is simply

$$\frac{d\phi}{dt} = \nabla\phi \cdot d\mathbf{r} \tag{2.42}$$

Setting $\mathbf{v} = \frac{d\mathbf{r}}{dt}$ we obtain from the definition of the scalar product

$$\frac{d\phi}{dt} = |\nabla\phi||\mathbf{v}| \cos(\theta) \tag{2.43}$$

where θ is the angle between $\nabla\phi$ and $\mathbf{v}$. This also shows that the largest increase of the scalar field is obtained in the direction $\nabla\phi$ and that, if $\mathbf{v}$ is a unit vector the rate of change in this case equals $|\nabla\phi|$.

We can extend the above analysis to find the rate of change of a vector field in a particular direction $\mathbf{v} = \frac{d\mathbf{r}}{dt}$.

$$\begin{array}{ll} \frac{d}{dt}\mathbf{a}(\mathbf{r}(t)) & = \sum_i \frac{da_i(\mathbf{r}(t))}{dt}\mathbf{e}_i \\ & = \sum_i \left(\sum_j (\partial_j a_i)(\mathbf{r}))v_j\right)\mathbf{e}_i \\ & = \sum_j v_j \partial_j \left(\sum_i a_i(\mathbf{r})\mathbf{e}_i\right) \\ & = \sum_j v_j \partial_j \mathbf{a}(\mathbf{r}) \\ & = (\mathbf{v}\cdot\nabla)\,\mathbf{a}(\mathbf{r}) \end{array} \tag{2.44}$$

This shows that the operator

$$\mathbf{v}\cdot\nabla = \sum_i v_i \partial_i \tag{2.45}$$

gives the rate of change in direction of $\mathbf{v}$ of the quantity (vector or scalar) on which it acts.

A second interesting geometrical property of $\nabla\phi$ may be found by considering the surface defined by $\phi(\mathbf{r}) = c$, where c is some constant. If $\mathbf{r}(t)$ is a space curve in the surface we clearly have $\frac{d\phi}{dt} = 0$ and consequently for any tangent vector $\mathbf{v}$ in the plane we have $\nabla\phi \cdot \mathbf{v} = 0$, according to Eq. (2.42). In other words, $\nabla\phi$ is a vector *normal* to the surface $\phi(\mathbf{r}) = c$ at every point.

Divergence of a vector field

The *divergence* of a vector field $\mathbf{a}(\mathbf{r})$ is defined by

$$(\text{div } \mathbf{a})(\mathbf{r}) = (\nabla\cdot\mathbf{a})(\mathbf{r}) = \sum_i (\partial_i a_i)(\mathbf{r}) \tag{2.46}$$

The full physical and geometrical meaning of the divergence is discussed in next section. Clearly $(\nabla \cdot \mathbf{a})(\mathbf{r})$ is a scalar field. Now if some vector field $\mathbf{a}$ is itself derived from a scalar field via $\mathbf{a} = \nabla\phi$ then $\nabla \cdot \mathbf{a}$ has the form $\nabla \cdot \nabla\phi$ or, as it is usually written $\nabla^2\phi$ where

$$\nabla^2 = \sum_i \partial_{ii}^2 \qquad \Big(= \sum_i \frac{\partial^2}{\partial x_i^2} \Big)$$

$\nabla^2\phi$ is called the *laplacian* of ϕ and is typically encountered in electrostatic problems or in diffusion equations of physical scalar field such as a temperature or density distribution. The laplacian can also act on vector through the components

$$\nabla^2\mathbf{a} = \sum_i \partial_{ii}^2\mathbf{a} = \sum_j \sum_i (\partial_{ii}^2 a_j)\mathbf{e}_j$$

Curl of a vector field

The *curl* of a vector field $\mathbf{a}(\mathbf{r})$ is defined by

$$\text{curl }(\mathbf{a}) = \nabla \times \mathbf{a} = \big(\partial_2 a_3 - \partial_3 a_2\big)\mathbf{e}_1 + \big(\partial_3 a_1 - \partial_1 a_3\big)\mathbf{e}_2 + \big(\partial_1 a_2 - \partial_2 a_1\big)\mathbf{e}_3,$$

In analogy to the definition of cross-product between two ordinary vectors we can express the definition symbolically

$$\nabla \times \mathbf{a} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ a_1 & a_2 & a_3 \end{pmatrix},$$

where it is understood that, on expanding the determinant, the partial derivatives act on the components of a. The tensor notation of this expression (for the i'th component) is even more compact

$$[\nabla \times \mathbf{a}]_i = \sum_{jk} \epsilon_{ijk}\partial_j a_k \tag{2.47}$$

Clearly, $\nabla \times \mathbf{a}$ is itself a vector field.

For a vector field $\mathbf{v}(\underline{x})$ describing the local velocity at any point in a fluid, $\nabla \times \mathbf{v}$ is a measure of the angular velocity of the fluid in the neighbourhood of that point. If a small paddle wheel were placed at various points in the fluid then it would tend to rotate in regfions where $\nabla \times \mathbf{v} \neq 0$, while it would not rotate in regions where $\nabla \times \mathbf{v} = 0$.

Another insight into the physical interpretation of the curl operator is gained by considering the vector field $\mathbf{v}$ describing the velocity at any point in a regied body rotation about some axis with angular velocity ω. If $\mathbf{r}$ is the position vector of the point with respect to some origin on the axis of rotation then the velocity field would be given by $\mathbf{v}(\underline{x}) = \omega \times \mathbf{r}(\underline{x})$. The curl of the vector field is then found to be $\nabla \times \mathbf{v} = 2\omega$.

Combinations of grad,div and curl

There are myriad of identities between various combinations of the three important vector operators grad, div and curl. The identities involving cross products are more easily proven using tensor calculus which is postponed for chapter 4. Here we simply list the most important identities:

$$\begin{array}{ll} \nabla(\phi+\psi) & = \nabla\phi + \nabla\psi \\ \nabla\cdot(\mathbf{a}+\mathbf{b}) & = \nabla\cdot\mathbf{a} + \nabla\cdot\mathbf{b} \\ \nabla\times(\mathbf{a}+\mathbf{b}) & = \nabla\times\mathbf{a} + \nabla\times\mathbf{b} \\ \nabla\cdot(\phi\mathbf{a}) & = \phi\nabla\cdot\mathbf{a} + \mathbf{a}\cdot\nabla\phi \\ \nabla\cdot(\mathbf{a}\times\mathbf{b}) & = \mathbf{b}\cdot(\nabla\times\mathbf{a}) - \mathbf{a}\cdot(\nabla\times\mathbf{b}) \\ \nabla\times(\nabla\times\mathbf{a}) & = \nabla(\nabla\cdot\mathbf{a}) - \nabla^2\mathbf{a} \\ \nabla\cdot(\nabla\times\mathbf{a}) & = 0 \\ \nabla\times\nabla\phi & = 0 \end{array} \tag{2.48}$$

Here, ϕ and ψ are scalar fields and $\mathbf{a}$ and $\mathbf{c}$ are vector fields. The last identity has an important meaning. If $\mathbf{a}$ is derived from the gradient of some scalar field, $\mathbf{a} = \nabla\phi$ then the identity shows that $\mathbf{a}$ is necessarily irrotational ,$\nabla\times\mathbf{a} = 0$. We shall return to this point in the next section.

2.6.5 Line, surface and volume integrals

In the previous section we have discussed continuously varying scalar and vector fields and discussed the action of various differential operators on the. Often, the need arises to consider the integration of field quantities along lines, over surfaces and throughout volumes. In general the integrand may be scalar or vector, but the evaluation of such intregrals involves their reduction to one or more scalar integrals, which are then evaluated. This procedure is equivalent to the way one in practice operates with differential operators on scalar and vector fields.

Line integrals

In general, one may encounter line integrals of the forms

$$\int_C \phi d\mathbf{r}, \qquad \int_C \mathbf{a}\cdot d\mathbf{r}, \qquad \int_C \mathbf{a}\times d\mathbf{r}, \tag{2.49}$$

where ϕ is a scalar field, $\mathbf{a}$ is a vector field and C is a prescribed curve in space joining to given points A and B in space. The three integrals themselves are respectively vector, scalar and vector in nature.

The formal definition of a line integral closely follow that of ordinary integrals and can be considered as the limit of a sum. We may divide the path joining the points A and B into N small line elements $\Delta\mathbf{r}_p$, $p = 1,\cdots,N$. If $\underline{x}_p = (x_{1,p}, x_{2,p}, x_{3,p})$ is any point on the line element $\Delta\mathbf{r}_p$ then the second type of

line integral in Eq. (2.49), for example, is defined as

$$\int_C \mathbf{a}\cdot d\mathbf{r} =_{\text{def}} \lim_{N\to\infty}\sum_{p=}^{N}\mathbf{a}(\underline{x}_p)\cdot\Delta\mathbf{r}_p,$$

where all $|\Delta\mathbf{r}_p|\to 0$ as $N\to\infty$.

Each of the line integrals in Eq. (2.49) is evaluated over some curve C that may be either open (A and B being distinct points) or closed (A and B coincide). In the latter case, one writes $\oint_C$ to indicate this. The curve C is a space curve, c.f. section 2.5.1, most often defined in a parametric form. In a cartesian coordinate system $\mathcal{C}=(O,\mathcal{R})$, it becomes

$$C: \quad \mathbf{r}(u)=\sum_i x_i(u)\mathbf{e}_i, \qquad u_0\le u\le u_1, \quad \text{and} \quad [\mathbf{r}(u_0)]=\underline{x}(A),\ [\mathbf{r}(u_1)]=\underline{x}(B) \tag{2.50}$$

The metod of evaluating a line integral is to reduce it to a set of scalar integrals. It is usual to work in cartesian coordinates, in which case, $d\mathbf{r}=\sum_i \mathbf{e}_i dx_i$. The three integrals in Eq. (2.49) then becomes

$$\begin{array}{lll} \int_C \phi(\mathbf{r})d\mathbf{r} & = \int_C \phi(\mathbf{r})\left(\sum_i \mathbf{e}_i dx_i\right) & = \sum_i \left(\int_C \phi(\mathbf{r})dx_i\right)\mathbf{e}_i \\ \int_C \mathbf{a}(\mathbf{r})\cdot d\mathbf{r} & = \int_C \left(\sum_i a_i(\mathbf{r})\mathbf{e}_i\right)\cdot\left(\sum_j \mathbf{e}_j dx_j\right) & = \sum_i \int_C a_i(\mathbf{r})dx_i \end{array} \tag{2.51}$$

and

$$\begin{array}{rcl} \int_C \mathbf{a}(\mathbf{r})\times d\mathbf{r} & = & \int_C \left(\sum_i a_i(\mathbf{r})\right)\times\left(\sum_j \mathbf{e}_j dx_j\right) \\ & = & \left(\int_C a_2(\mathbf{r})dx_3-\int_C a_3(\mathbf{r})dx_2\right)\mathbf{e}_1+\left(\int_C a_3(\mathbf{r})dx_1-\int_C a_1(\mathbf{r})dx_3\right)\mathbf{e}_2 \\ & & +\left(\int_C a_1(\mathbf{r})dx_2-\int_C a_2(\mathbf{r})dx_1\right)\mathbf{e}_3 \end{array} \tag{2.52}$$

Note, that in the above we have used relations of the form

$$\int_C a_i\mathbf{e}_j dx_j=\left(\int_C a_i dx_j\right)\mathbf{e}_j,$$

which is allowable since the cartesian basis is independent of the coordinates. If $\mathcal{E}=\mathcal{E}(\underline{x})$ then the basis vectors could not be factorised out in the integral. The final scalar integrals in Eq. (2.51) and Eq. (2.52) can be solved using the parametrized form for C, Eq. (2.50). For instance

$$\int_C a_2(\mathbf{r})dx_3=\int_{u_0}^{u_1} a_2(x_1(u),x_2(u),x_3(u))\frac{dx_3}{du}du.$$

In general, a line integral between two points will depend on the specific path C defining the integral. However, for line integrals of the form $\int_C \mathbf{a}\cdot d\mathbf{r}$ there exists a class of vector fields for which the line integral between two points is *independent* of the path taken. Such vector fields are called *conservative*. A vector field $\mathbf{a}$ that has continuous partial derivatives in a simply connected region R [7], is conservative if, and only if, any of the following is true

[7] A simply connected region R is a region in space for which any closed path can be continuously shrunk to a point, ie. R has no "holes"

1. The integral $\int_A^B \mathbf{a} \cdot d\mathbf{r}$, where $A, B \in R$, is independent of the path from A to B. Hence $\oint_C \mathbf{a} \cdot d\mathbf{r} = 0$ around any closed loop in R.

2. There exists a scalar field ϕ in R such that $\mathbf{a} = \nabla\phi$.

3. $\nabla \times \mathbf{a} = 0$.

4. $\mathbf{a} \cdot d\mathbf{r}$ is an exact differential.

We will not demonstrate the equivalence of these statements. If a vector field is conservative we can write

$$\mathbf{a} \cdot d\mathbf{r} = \nabla\phi \cdot d\mathbf{r} = d\phi$$

and

$$\int_A^B \mathbf{a} \cdot d\mathbf{r} = \int_A^B \nabla\phi \cdot d\mathbf{r} = \int_A^B d\phi = \phi(A) - \phi(B).$$

This situtation is encountered whenever $\mathbf{a} = \mathbf{f}$ represents the force $\mathbf{f}$ derived from a potential (scalar) field, ϕ, such as the potential energy in a gravitational field, the potential energy in an elastic spring, the voltage in a electrical circuits, etc.

Surface integrals

As with line integrals, integrals over surfaces can involve vector and scalar fields and, equally, result in either a vector or a scalar. We shall focus on surface integrals of the form

$$\int_S \mathbf{a} \cdot d\mathbf{S}. \tag{2.53}$$

where $\mathbf{a}$ is a vector field and S is a surface in space which may be either open or closed. Following the notation of line integrals, for surface integrals over a closed surface $\int_S$ is replaced by $\oint_S$. The vector differential $d\mathbf{S}$ in Eq. (2.53) represents a vector area element of the surface S. It may also be written $d\mathbf{S} = \mathbf{n}dS$ where $\mathbf{n}$ is a unit normal to the surface at the position of the element and dS is a scalar area of the element. The convention for the direction of the normal $\mathbf{n}$ to a surface depends on whether the suface is open or closed. For a closed surface the direction of $\mathbf{n}$ is taken be outwards from the enclosed volume. An open surface spans some perimeter curve C. The direction of $\mathbf{n}$ is then given by the right-hand sense with respect to the direction in which the perimeter is traversed, ie. it follows the right-hand screw rule.

The formal definition of a surface integral is very similar to that of a line integral. One divides the surface into N elements of area ΔS_p, $p = 1, \cdots, N$ each with a unit normal $\mathbf{n}_p$. If $\underline{x}_p$ is any point in ΔS_p then

$$\int_S \mathbf{a} \cdot d\mathbf{S} = \lim_{N\to\infty} \sum_{p=1}^{N} \mathbf{a}(\underline{x}_p) \cdot \mathbf{n}_p \Delta S_p,$$

where it is required that $\Delta S_p \to 0$ for $N \to \infty$.

A standard way of evaluating surface integrals is to use cartesian coordinates and project the surface onto one of the basis planes. For instance, suppose a surface S has projection R onto the 12-plane (xy-plane), so that an element of surface area dS projects onto the area element dA. Then

$$dA = |\mathbf{e}_3 \cdot d\mathbf{S}| = |\mathbf{e}_3 \cdot \mathbf{n} dS| = |\mathbf{e}_3 \cdot \mathbf{n}| dS.$$

Since in the 12-plane $dA = dx_1 dx_2$ we have the expression for the surface integral

$$\int_S \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} = \int_R \mathbf{a}(\mathbf{r}) \cdot \mathbf{n} dS = \int_R \mathbf{a}(\mathbf{r}) \cdot \mathbf{n} \frac{dx_1 dx_2}{|\mathbf{e}_3 \cdot \mathbf{n}|}$$

Now, if the surface S is given by the equation $x_3 = z(x_1, x_2)$, where $z(x_1, x_2)$ gives the third coordinate of the surface for each (x_1, x_2) then the scalar field

$$f(x_1, x_2, x_3) = x_3 - z(x_1, x_2) \tag{2.54}$$

is identical zero on S. The unit normal at any point of the surface will be given by $\mathbf{n} = \frac{\nabla f}{|\nabla f|}$ evaluated at that point, c.f. section 2.6.4. We then obtain

$$d\mathbf{S} = \mathbf{n} dS = \frac{\nabla f}{|\nabla f|} \frac{dA}{|\mathbf{n} \cdot \mathbf{e}_3|} = \nabla f \frac{dA}{|\nabla f \cdot \mathbf{e}_3|} = \nabla f \frac{dA}{|\partial_3 f|} = \nabla f \; dx_1 dx_2,$$

where the last identity follows from the fact that $\partial_3 f = 1$ from Eq. (2.54). The surface integral then becomes

$$\int_S \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} = \int_R \mathbf{a}(x_1, x_2, z(x_1, x_2)) \cdot (\nabla f)(x_1, x_2) dx_1 dx_2.$$

which is an ordinary scalar integral in two variables x_1 and x_2.

Volume integrals

Volume integrals are generally simpler than line or surface integrals since the element of the volume dV is a scalar quantity. Volume integrals are most often on the form

$$\int_V \phi(\mathbf{r}) dV \qquad \int_V \mathbf{a}(\mathbf{r}) dV \tag{2.55}$$

Clearly, the firs form results in a scalar, whereas the second one yields a vector. Two closely related physical examples, one of each kind, are provided by the total mass M, of a fluid contained in a volume V, given by $M = \int_V \rho(\mathbf{r}) dV$ and the total linear momentum of that same fluid, given by $\int_V \rho(\mathbf{r}) \mathbf{v}(\mathbf{r}) dV$, where $\rho(\mathbf{r})$ is the density field and $\mathbf{v}(\mathbf{r})$ is the velocity field of the fluid.

The evaluation of the first volume integral in Eq. (2.55) is an ordinary multiple integral. The evaluation of the second type of volume integral follows directly since we can resolve the vector field into cartesian coordinates

$$\int_V \mathbf{a}(\mathbf{r}) dV = \sum_i \Big(\int_V a_i(\mathbf{r}) dV \Big) \mathbf{e}_i.$$

Of course we could have written $\mathbf{a}$ in terms of the basis vectors of other coordinate system (e.g. spherical coordinates) but since such basis vectors are not in general constant, they cannot be taken out of the integrand.

2.6.6 Integral theorems

There are two important theorems relating surface and volume integrals of vector fields, the *divergence theorem* of Gauss and Stokes theorem.

Gauss theorem

Let $\mathbf{a}$ be a (differentiable) vector field, and V be any volume bounded by a closed surface S. Then Gauss theorem states

$$\int_V \big(\nabla \cdot \mathbf{a}\big)(\mathbf{r}) dV = \oint_S \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S}. \tag{2.56}$$

, The proof goes as follows. Let $\mathcal{C} = (O, \mathcal{R})$ be a cartesian coordinate system, and let $B(\mathbf{r}_0)$ be a small box around $\mathbf{r}_0$ with its edges oriented along the directions of the basis vectors and with volume $V_B = \Delta x_1 \Delta x_2 \Delta x_3$. Each face can be labelled as sk where $s = \pm 1$ and $k = 1, 2, 3$ indicates the orientation of the face. Thus the outward normal of face sk is $\mathbf{n}_{sk} = s\mathbf{e}_k$. If the area $A_k = \prod_{i \neq k} \Delta x_i$ of each face F_{sk} is small then we can approximate each surface integral

$$\int_{F_{sk}} \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} = \int_{F_{sk}} \mathbf{a}(\mathbf{r}) \cdot s\mathbf{e}_k dS$$

by evaluating the vector field $\mathbf{a}(\mathbf{r})$ in the center point of the face $\mathbf{r}_0 + s\frac{\Delta x_k}{2}\mathbf{e}_k$

$$\int_{F_{sk}} \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} \approx \mathbf{a}\left(\mathbf{r}_0 + s\frac{\Delta x_k}{2}\mathbf{e}_k\right) \cdot s\mathbf{e}_k \; A_k$$

The surface integral on the rhs. of Eq. (2.56) for the box $B(\mathbf{r}_0)$ then becomes

$$\begin{aligned}
\oint_{B(\mathbf{r}_0)} \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} \quad &\approx \textstyle\sum_{sk} \mathbf{a}\left(\mathbf{r}_0 + s\frac{\Delta x_k}{2}\mathbf{e}_k\right) \cdot s\mathbf{e}_k \; A_k \\
&= \textstyle\sum_k \Big(\mathbf{a}\left(\mathbf{r}_0 + \frac{\Delta x_k}{2}\mathbf{e}_k\right) - \mathbf{a}\left(\mathbf{r}_0 - \frac{\Delta x_k}{2}\mathbf{e}_k\right)\Big) \cdot \mathbf{e}_k \; A_k \\
&= \textstyle\sum_k \Big(a_k\left(\mathbf{r}_0 + \frac{\Delta x_k}{2}\mathbf{e}_k\right) - a_k\left(\mathbf{r}_0 - \frac{\Delta x_k}{2}\mathbf{e}_k\right)\Big) \; A_k \\
&\approx \textstyle\sum_k (\partial_k a_k)(\mathbf{r}_0) \; \Delta x_k A_k \\
&= \big(\nabla \cdot \mathbf{a}\big)(\mathbf{r}_0) \; V_B
\end{aligned} \tag{2.57}$$

Any volume V can be approximated by a set of small boxes, B_i, centered around the points $\mathbf{r}_{0,i}$, where $i = 1, \cdots, N$. For the lhs. of Eq. (2.56) we then have

$$\int_V \nabla \cdot \mathbf{a}(\mathbf{r}) dV \approx \sum_i \big(\nabla \cdot \mathbf{a}\big)(\mathbf{r}_{0,i}) \; V_{B_i}. \tag{2.58}$$

Adding the surface integrals of each of these boxes, the contribution from the mutual interfaces vanishes (since the outward normal of the two adjacent boxes

point in opposite direction). Consequently, the only contributions comes from the surface S of V.

$$\oint_S \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} \approx \sum_i \oint_{B(\mathbf{r}_{0,i})} \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S}_i \tag{2.59}$$

Since each term on the rhs. of Eq. (2.58) equals a corresponding term on the rhs. of Eq. (2.59) the Gauss teorem is demonstrated.

Gauss theorem is often used in conjunction with following mathematical theorem

$$\frac{d}{dt}\int_V \phi(\mathbf{r},t)dV = \int_V \frac{\partial}{\partial t}\phi(\mathbf{r},t)dV,$$

where t is time and ϕ is a time dependent scalar field (The theorem works in arbitrary spatial dimension). The two theorems are central in deriving partial differential equations for dynamical systems, in particular the so called *continuity equations*, linking a flow field to the time changes of scalar field advected by the flow. For instance if $\rho(\mathbf{r},t)$ is a density and $\mathbf{v}(\mathbf{r},t)$, is the velocity field of a fluid then the vector $\mathbf{j}(\mathbf{r}) = \rho(\mathbf{r})\mathbf{v}(\mathbf{r})$ gives the density current. It can then by shown (try it) that under mass conservation then

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0.$$

The divergence of a vector field therefore has the physical meaning of giving the net "outflux" of a scalar advected by the field within an infinitesimal volume.

Stokes theorem

Stokes theorem states that if S is the "curl analogue" of the divergence theorem and relates the integral of the curl of a vector field over an open surface S to the line integral of the vector field around the perimeter C bounding the surface.

$$\int_S \big(\nabla \times \mathbf{a}\big)(\mathbf{r}) \cdot d\mathbf{S} = \oint_C \mathbf{a}(\mathbf{r}) \cdot d\mathbf{r} \tag{2.60}$$

Following the same lines as for the derivation of the divergence theorem the surface S can be divided into many small areas S_i with boundaries C_i and unit normals $\mathbf{n}_i$. For each small area one can show that

$$(\nabla \times \mathbf{a}) \cdot \mathbf{n}_i S_i \approx \oint_{C_i} \mathbf{a} \cdot d\mathbf{r}.$$

Summing over i one finds that on the rhs. all parts of all interior boundaries that are not part of C are included twice, being traversed in opposite directions on each occasion and thus cancelling each other. Only contributions from line elements that are also part of C survive. If each S_i is allows to to tend to zero, Stokes theorem, Eq. (2.60), is obtained.

2.7 Curvilinear coordinates

The vector operators, $\nabla\phi$, $\nabla \cdot \mathbf{a}$ and $\nabla \times \mathbf{a}$, we have discussed so far have all been defined in terms of cartesian coordinates. In that respect we have been more restricted than in the algebraic definitions of the analogous ordinary scalar and cross product, Eq. (2.18) and Eq. (2.20) respectively. Here we only assumed orthonormality of the basis. The reason is that the nabla operator involves spatial derivatives which implies that one must account for the possible non-constancy of $\mathcal{E}$, c.f. Eq. (2.35).

Many systems possess some particular symmetry which makes other coordinate systems more natural, notably cylindrical or spherical coordinates. These coordinates are just two examples of what are called *curvilinear coordinates*. Curvilinear coordinates refer to the general case in which the rectangular coordinates $\underline{x}$ of any point can be expressed as functions of another set of coordinates $\underline{x}'$, thus defining a coordinate transformation, $\underline{x} = \underline{x}(\underline{x}')$ or $\underline{x}' = \underline{x}'(\underline{x})$ [8]. Here we shall discuss the algebraic form of the standard vector operators for the two particular transformations into cylindrical or spherical coordinates. An important feature of these coordinates is that although they lead to non-constant bases, these bases retain the property of orthonormality.

The starting point for the disgression is to realize that a non-constant basis

$$\mathcal{E}'(\underline{x}') = \{\mathbf{e}'_1(\underline{x}'), \mathbf{e}'_2(\underline{x}'), \mathbf{e}'_3(\underline{x}')\},$$

implies that the basis vectors are *vector fields*, as opposed to the constant basis $\mathcal{R} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ associated with cartesian coordinates. Any vector field, $\mathbf{a}(\underline{x}')$ is then generally written as

$$\mathbf{a}(\underline{x}') = \sum_j a'_j(\underline{x}')\mathbf{e}'_j(\underline{x}') \tag{2.61}$$

where a'_j are the components of the vector field in the new basis, and $\underline{x}' = \underline{x}'(\underline{x})$ is the coordinate transformation. One notes, that the functional form of the components $a'_j(\underline{x}')$ differ from the functional form of the cartesian components, $a_j(\underline{x})$, because the same point, P will have two different numerical representations $\underline{x}'(P)$ and $\underline{x}(P)$. For a scalar field one must have the identity

$$\phi'(\underline{x}') = \phi(\underline{x}),$$

where ϕ' is the functional form of the scalar field in the primed coordinate system and ϕ is the function of the scalar field with respect to the unprimed system. Again, the two functional forms must be different because the same point has different numerical representations. However, the value of the two functions must be the same since $\underline{x}'$ and $\underline{x}$ represent the same point in space.

[8] Recall that the notation $\underline{x}' = \underline{x}'(\underline{x})$ is a short hand notation for the three functional relationships listed in Eq. (1.2).

2.7.1 Cylindrical coordinates

Cylindrical coordinates, $\underline{x}' = (\rho, \phi, z)$, are defined in terms of normal cartesian coordinates $\underline{x} = (x_1, x_2, x_3) = (x, y, z)$ by the coordinate transformation

$$\begin{aligned} x &= \rho\cos(\phi) \\ y &= \rho\sin(\phi), \\ z &= z, \end{aligned} \tag{2.62}$$

The domain of variation is $0 \leq \rho < \infty, \quad 0 \leq \phi < 2\pi$.

The position vector may be written as

$$\mathbf{r}(\underline{x}') = \rho\cos(\phi)\mathbf{e}_1 + \rho\sin(\phi)\mathbf{e}_2 + z\mathbf{e}_3$$

Local basis

If we take the partial derivatives with respect to ρ, ϕ, z, c.f. Eq. (2.33), and normalize we obtain

$$\begin{array}{llll} \mathbf{e}'_1(\underline{x}') & = \mathbf{e}_\rho(\underline{x}') & = \partial_\rho \mathbf{r} & = \cos(\phi)\mathbf{e}_1 + \sin(\phi)\mathbf{e}_2 \\ \mathbf{e}'_2(\underline{x}') & = \mathbf{e}_\phi(\underline{x}') & = \frac{1}{\rho}\partial_\phi \mathbf{r} & = -\sin(\phi)\mathbf{e}_1 + \cos(\phi)\mathbf{e}_2 \\ \mathbf{e}'_3 & = \mathbf{e}_3 & & \end{array}$$

These three unit vectors, like the Cartesian unit vectors $\mathbf{e}_i$, form an orthonormal basis at each point in space. An arbitrary vector field may therefore be resolved in this basis

$$\mathbf{a}(\underline{x}') = a_r(\underline{x}')\mathbf{e}_r(\underline{x}') + a_\phi(\underline{x}')\mathbf{e}_\phi(\underline{x}') + a_z(\underline{x}')\mathbf{e}_z$$

where the vector components

$$a_r = \mathbf{a}\cdot\mathbf{e}_r, \quad a_\phi = \mathbf{a}\cdot\mathbf{e}_\phi, \quad V_z = \mathbf{a}\cdot\mathbf{e}_z$$

are the projections of $\mathbf{a}$ on the local basis vectors.

Resolution of gradient

The derivatives after cylindrical coordinates are found by differentiation through the Cartesian coordinates (chain rule)

$$\begin{array}{llll} \partial_\rho & = \frac{\partial x}{\partial \rho}\partial_x + \frac{\partial y}{\partial \rho}\partial_y & = \cos(\phi)\partial_x + \sin(\phi)\partial_y \\ \partial_\phi & = \frac{\partial x}{\partial \phi}\partial_x + \frac{\partial y}{\partial \phi}\partial_y & = -\rho\sin(\phi)\partial_x + \rho\cos(\phi)\partial_y \end{array}$$

From these relations we can calculate the projections of the gradient operator $\nabla = \sum_i \mathbf{e}_i\partial_i$ on the cylindrical basis and we obtain

$$\begin{array}{llll} \nabla_\rho & = \mathbf{e}_\rho\cdot\nabla & = \partial_\rho \\ \nabla_\phi & = \mathbf{e}_\phi\cdot\nabla & = \frac{1}{\rho}\partial_\phi \\ \nabla_z & = \mathbf{e}_z\cdot\nabla & = \partial_z \end{array}$$

The resolution of the gradient in the two bases therefore becomes

$$\nabla = \mathbf{e}_\rho \nabla_\rho + \mathbf{e}_\phi \nabla_\phi + \mathbf{e}_z \nabla_z = \mathbf{e}_\rho \partial_\rho + \mathbf{e}_\phi \frac{1}{\rho} \partial_\phi + \mathbf{e}_z \partial_z$$

Together with the only non-vanishing derivatives of the basis vectors

$$\begin{aligned} \partial_\phi \mathbf{e}_\rho &= \mathbf{e}_\phi \\ \partial_\phi \mathbf{e}_\phi &= -\mathbf{e}_\rho \end{aligned}$$

we have the necessary tools for calculating in cylindrical coordinates. Note, that it is the vanishing derivatives of the basis vectors that lead to the simple form of the vector operators in cartesian coordinates.

Laplacian

The laplacian in cylindrical coordinates takes the form

$$\nabla^2 = \nabla \cdot \nabla = (\mathbf{e}_\rho \nabla_\rho + \mathbf{e}_\phi \nabla_\phi + \mathbf{e}_z \nabla_z) \cdot (\mathbf{e}_\rho \nabla_\rho + \mathbf{e}_\phi \nabla_\phi + \mathbf{e}_z \nabla_z)$$

Using the linearity of the scalar product it can be rewritten to the form

$$\sum \mathbf{e}'_i \nabla'_i \cdot \mathbf{e}'_j \nabla'_j,$$

where $\mathbf{e}'_1 = \mathbf{e}_\rho$, $\nabla'_1 = \nabla_\rho$ etc. The interpretation of each of these terms is

$$\mathbf{e}'_i \nabla'_i \cdot \mathbf{e}'_j \nabla'_j = \mathbf{e}'_i \cdot \nabla'_i (\mathbf{e}'_j \nabla'_j) \tag{2.63}$$

Applying the chain rule of differentiation and Eq. (2.7.1) one can then show

$$\nabla^2 = \partial^2_{\rho\rho} + \frac{1}{\rho} \partial_\rho + \frac{1}{\rho^2} \partial^2_{\phi\phi} + \partial^2_{zz}$$

2.7.2 Spherical coordinates

The treatment of spherical coordinates follow the same line as cylindrical coordinates. The spherical coordinates are $\underline{x}' = (r, \phi, \theta)$ and the coordinate transformation is given by

$$\begin{aligned} x &= r \sin(\theta) \cos(\phi) \\ y &= r \sin(\theta) \sin(\phi) \\ z &= r \cos(\theta) \end{aligned} \tag{2.64}$$

The domain of variation is $0 \leq r < \infty, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi$. The position vector is

$$\mathbf{r}(\underline{x}') = r \sin(\theta) \cos(\phi) \mathbf{e}_1 + r \sin(\theta) \sin(\phi) \mathbf{e}_2 + r \cos(\theta) \mathbf{e}_3$$

Local basis

The normalized tangent vectors along the directions of the spherical coordinates are

$$\begin{array}{lll} \mathbf{e}_r(\underline{x}') & = \partial_r \mathbf{r} & = \sin(\theta)\cos(\phi)\mathbf{e}_1 + \sin(\theta)\sin(\phi)\mathbf{e}_2 - \sin(\theta)\mathbf{e}_3 \\ \mathbf{e}_\theta(\underline{x}') & = \frac{1}{r}\partial_\theta \mathbf{r} & = \cos(\theta)\cos(\phi)\mathbf{e}_1 + \cos(\theta)\sin(\phi)\mathbf{e}_2 - \sin(\theta)\mathbf{e}_3 \\ \mathbf{e}_\phi(\underline{x}') & = \frac{1}{r\sin(\theta)}\partial_\phi \mathbf{r} & = -\sin(\phi)\mathbf{e}_1 + \cos(\phi)\mathbf{e}_2 \end{array}$$

They define an orthonormal basis such that an arbitrary vector field may be resolved in these directions

$$\mathbf{a}(\underline{x}') = a_r(\underline{x}')\mathbf{e}_r(\underline{x}') + a_\theta(\underline{x}')\mathbf{e}_\theta(\underline{x}') + a_\phi(\underline{x}')\mathbf{e}_\phi(\underline{x}')$$

with $a_i' = \mathbf{e}_i' \cdot \mathbf{a}$, where $a_1' = a_r$, $\mathbf{e}_1' = \mathbf{e}_r$ etc.

Resolution of the gradient

The gradient operator may also be resolved on the basis

$$\begin{array}{ll} \nabla_r & = \mathbf{e}_r\nabla_r + \mathbf{e}_\theta\nabla_\theta + \mathbf{e}_\phi\nabla_\phi \\ & = \mathbf{e}_r\partial_r + \mathbf{e}_\theta\frac{1}{r}\partial_\theta + \frac{1}{r\sin(\theta)}\partial_\phi \end{array} \tag{2.65}$$

The non vanishing derivatives of the basis vectors are

$$\begin{array}{llll} \partial_\theta \mathbf{e}_r & = \mathbf{e}_\theta & \partial_\phi \mathbf{e}_r & = \sin(\theta)\mathbf{e}_\phi \\ \partial_\theta \mathbf{e}_\theta & = -\mathbf{e}_r & \partial_\phi \mathbf{e}_\theta & = \cos(\theta)\mathbf{e}_\phi \\ & & \partial_\phi \mathbf{e}_\phi & = -\sin(\theta)\mathbf{e}_r - \cos(\theta)\mathbf{e}_\theta \end{array} \tag{2.66}$$

This is all we need for expressing vector operators in spherical coordinates.

Laplacian

The laplacian in spherical coordinates becomes

$$\nabla^2 = (\mathbf{e}_r\nabla_r + \mathbf{e}_\theta\nabla_\theta + \mathbf{e}_\phi\nabla_\phi)\cdot(\mathbf{e}_r\nabla_r + \mathbf{e}_\theta\nabla_\theta + \mathbf{e}_\phi\nabla_\phi)\,,$$

which, after using the linearity of the scalar product and Eq. (2.66) becomes

$$\nabla^2 = \partial_{rr}^2 + \frac{2}{r}\partial_r + \frac{1}{r^2}\partial_{\theta\theta}^2 + \frac{\cos(\theta)}{r^2\sin(\theta)}\partial_\theta + \frac{1}{r^2\sin^2(\theta)}\partial_{\phi\phi}^2$$

Notice that the two first terms involving radial derivatives can be given alternative expressions

$$(\partial_{rr}^2\phi) + (\frac{2}{r}\partial_r\phi) = \frac{1}{r^2}\Big(\partial_r\big(r^2(\partial_r\phi)\big)\Big) = \frac{1}{r}\big(\partial_{rr}^2(r\phi)\big),$$

where ϕ is a scalar field.

Chapter 3

Tensors

In this chapter we will limit ourself to the discussion of rank 2 tensors, unless stated otherwise. The precise definition of the rank of a tensor will become clear later.

3.1 Definition

A tensor, $\mathbf{T}$, of rank two is a geometrical object that to any vector $\mathbf{a}$ associate another vector $\mathbf{u} = \mathbf{T}(\mathbf{a})$ by a *linear* operation

$$\mathbf{T}(m\mathbf{a}) = m\mathbf{T}(\mathbf{a}) \tag{3.1}$$

$$\mathbf{T}(\mathbf{a}+\mathbf{b}) = \mathbf{T}(\mathbf{a}) + \mathbf{T}(\mathbf{b}) \tag{3.2}$$

In other words a rank 2 tensor is a *linear vector operator*. We will denote tensors (of rank 2 or higher) with capital bold-face letters.

Any linear transformation of vectors, such as rotations, reflections or projections are examples of tensors. In fact, we have already encountered tensors in disguise. The operator $\mathbf{T} = \mathbf{c}\times$ is a tensor. For each vector, $\mathbf{a}$, it associates the vector $\mathbf{T}(\mathbf{a}) = \mathbf{c} \times \mathbf{a}$, obtained by rotating $\mathbf{a}$ 90^0 counter-clockwise around $\mathbf{c}$ and scaling it with the magnitude $|\mathbf{c}|$. Since the cross-product is linear in the second argument $\mathbf{T}$ is linear and thus a tensor. For reasons to become clear we will also extend and use the dot-notation to indicate the operation of a tensor on a vector so

$$\mathbf{T} \cdot \mathbf{a} =_{\text{def.}} \mathbf{T}(\mathbf{a}) \tag{3.3}$$

Physical examples of tensors include the *inertia tensor*, $\mathbf{I}$, or the moment of inertia, that specifies how a torque, $\boldsymbol{\tau}$ (a vector), changes the angular momentum, $\frac{d\mathbf{l}}{dt}$ (a vector)

$$\frac{d\mathbf{l}}{dt} = \mathbf{I} \cdot \boldsymbol{\tau}$$

Another example is the (transposed) *stress* tensor, σ^t, that specifies the force $\mathbf{f}$ a continuous medium exerts on a surface element defined by the normal vector

n.

$$\mathbf{f} = \sigma^t \cdot \mathbf{n}$$

A tensor often associated with the stress tensor is the *strain* tensor, **U**, that specifies how a strained material is distorted $\Delta\mathbf{u}$ in some direction $\Delta\mathbf{r}$

$$\Delta\mathbf{u} = \mathbf{U} \cdot \Delta\mathbf{r}$$

We shall be more specific on these physical tensors later.

3.2 Outer product

The *outer product* $\mathbf{a} \otimes \mathbf{b}$ or simply **ab** between two vectors **a** and **b** is a tensor defined by the following equation, where **c** is any vector

$$(\mathbf{ab})(\mathbf{c}) = \mathbf{a}(\mathbf{b} \cdot \mathbf{c}) \tag{3.4}$$

In words, for each **c** the tensor **ab** associates a vector in the direction of **a** and with a magnitude equal to the projection of **c** into **b**. In order to call the object (**ab**) a tensor we should verify that it is a linear operator. Using the definition, Eq. (3.3), allows us to "place the brackets where we want"

$$(\mathbf{ab}) \cdot \mathbf{c} = \mathbf{a}(\mathbf{b} \cdot \mathbf{c}),$$

which will ease the notation. Now, demonstrating that (**ab**) indeed is a a tensor amounts to "moving brackets":

$$\begin{array}{rl} (\mathbf{ab}) \cdot (m\mathbf{c}) & = \mathbf{a}(\mathbf{b} \cdot (m\mathbf{c})) = m\mathbf{a}(\mathbf{b} \cdot \mathbf{c}) \\ & = m(\mathbf{ab}) \cdot \mathbf{c} \\ (\mathbf{ab}) \cdot (\mathbf{c} + \mathbf{d}) & = \mathbf{a}(\mathbf{b} \cdot (\mathbf{c} + \mathbf{d})) = \mathbf{a}(\mathbf{b} \cdot \mathbf{c}) + \mathbf{a}(\mathbf{b} \cdot \mathbf{d}) \\ & = (\mathbf{ab}) \cdot \mathbf{c} + (\mathbf{ab}) \cdot \mathbf{d} \end{array} \tag{3.5}$$

We note that since the definition, eq. (**??**) involves vector operations that bear no reference to coordinates/components, the outer product will itself be invariant to the choise of coordinate system. The outer product is also known in the litterature as the *tensor*, *direct*, *exterior* or *dyadic* product. The tensor formed by the outer product of two vectors is called a *dyad*.

3.3 Basic tensor algebra

We can form more general linear vector operators by taking sum of dyads. The sum of any two tensors,**S** and **T**, and the multiplication of a tensor with a scalar m are naturally defined by

$$\begin{array}{ll} (\mathbf{S} + \mathbf{T}) \cdot \mathbf{c} & = \mathbf{S} \cdot \mathbf{c} + \mathbf{T} \cdot \mathbf{c} \\ (m\mathbf{S}) \cdot \mathbf{c} & = m(\mathbf{S} \cdot \mathbf{c}). \end{array} \tag{3.6}$$

Here, $\mathbf{c}$ is any vector. It is easy to show that that $(\mathbf{S}+\mathbf{T})$ and $(m\mathbf{S})$ are also tensors, ie. they satisfy the definition of being linear vector operators, Eq.(4.3.2). Definition Eq. (3.6) and the properties Eq. (3.5) guarantee that dyad products, sums of tensors and dot products of tensors with vectors satisfy all the usual algebraic rules for sums and products. For instance, the outer product between two vectors on the form $\mathbf{c} = \sum_i m_i \mathbf{a}_i$ and $\mathbf{d} = \sum_j n_j \mathbf{b}_j$, where m_i and n_j are scalars, is

$$\mathbf{cd} = \left(\sum_i m_i \mathbf{a}_i\right)\left(\sum_j n_j \mathbf{b}_j\right) = \sum_{ij} m_i n_j \mathbf{a}_i \mathbf{b}_j, \tag{3.7}$$

ie. it is a sum of all dyad combinations $\mathbf{a}_i\mathbf{b}_j$. The sum of two or more dyads is called a *dyadic*. As we shall demonstrate in section 3.4 any tensor can be expressed as a dyadic but not necessarily as a single dyad.

We can also extend the definition of the dot product. For any vector $\mathbf{c}$ we define the dot product of two tensors by the following formula where $\mathbf{c}$ is any vector

$$(\mathbf{T}\cdot\mathbf{S})\cdot\mathbf{c} = \mathbf{T}\cdot(\mathbf{S}\cdot\mathbf{c}) \tag{3.8}$$

In words, application of the operator $(\mathbf{T}\cdot\mathbf{S})$ to any vector means first applying $\mathbf{S}$ and then $\mathbf{T}$. Since the association of two linear functions is a linear function $(\mathbf{T}\cdot\mathbf{S})$ is a tensor itself. Sums and products of tensors also obey usual rules of algebra except that dot multiplication of two tensors, in general, is not commutative

$$\mathbf{T}+\mathbf{S} = \mathbf{S}+\mathbf{T} \tag{3.9}$$

$$\mathbf{T}\cdot(\mathbf{S}+\mathbf{P}) = \mathbf{T}\cdot\mathbf{S}+\mathbf{T}\cdot\mathbf{P} \tag{3.10}$$

$$\mathbf{T}\cdot(\mathbf{S}\cdot\mathbf{P}) = (\mathbf{T}\cdot\mathbf{S})\cdot\mathbf{P} \tag{3.11}$$

3.3.1 Transposed tensors

Furthermore, we define a dot product of a dyad with a vector on the left in the obvious way

$$\mathbf{c}\cdot(\mathbf{ab}) = (\mathbf{c}\cdot\mathbf{a})\mathbf{b} \tag{3.12}$$

and correspondingly for a dyadic. Since the dot product of a dyadic with a vector is *not* in general commutative

$$\mathbf{T}\cdot\mathbf{c} \neq \mathbf{c}\cdot\mathbf{T}.$$

it makes sence to introduce the *transpose* $\mathbf{T}^t$ of a tensor

$$\mathbf{T}^t\cdot\mathbf{c} = \mathbf{c}\cdot\mathbf{T} \tag{3.13}$$

Since $\mathbf{c}\cdot(\mathbf{ab}) = (\mathbf{ba})\cdot\mathbf{c}$ we have for a dyad

$$(\mathbf{ab})^t = \mathbf{ba}. \tag{3.14}$$

As we shall see in section 3.4 any tensor can be expressed as a sum of dyads. Using this fact following properties can easily be proved

$$(\mathbf{T}+\mathbf{S})^t = \mathbf{T}^t + \mathbf{S}^t \tag{3.15}$$
$$(\mathbf{T}\cdot\mathbf{S})^t = \mathbf{S}^t \cdot \mathbf{T}^t \tag{3.16}$$
$$\left(\mathbf{T}^t\right)^t = \mathbf{T}. \tag{3.17}$$

3.3.2 Contraction

Another useful operation on tensors is that of *contraction*. The contraction, $\mathbf{ab} : \mathbf{cd}$, of two dyads results in a scalar defined by

$$\mathbf{ab} : \mathbf{cd} = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) \tag{3.18}$$

Note the following useful relation for the contraction of two dyads formed by basis vectors

$$\mathbf{e}_i\mathbf{e_j} : \mathbf{e}_k\mathbf{e}_l = \delta_{ik}\delta_{jl}$$

The contraction of two dyadics is defined as the bilinear operation

$$\left(\sum_i m_i\mathbf{a}_i\mathbf{b}_i\right) : \left(\sum_j n_j\mathbf{c}_j\mathbf{d}_j\right) = \sum_{ij} m_i n_j(\mathbf{a}_i\cdot\mathbf{c}_j)(\mathbf{b}_i\cdot\mathbf{d}_j), \tag{3.19}$$

where m_i and n_j are scalars.

In some textbook another type of contraction (a "transposed contraction") is also defined

$$\mathbf{ab}\cdot\cdot\mathbf{cd} = \mathbf{ab} : \mathbf{dc}$$

Similarly for two dyadics

$$\left(\sum_i m_i\mathbf{a}_i\mathbf{b}_i\right)\cdot\cdot\left(\sum_j n_j\mathbf{c}_j\mathbf{d}_j\right) = \sum_{ij} m_i n_j(\mathbf{a}_i\cdot\mathbf{d}_j)(\mathbf{b}_i\cdot\mathbf{c}_j)$$

3.3.3 Special tensors

Some special tensors arise as a consequence of the basic tensor algebra.
A *symmetric tensor* is a tensor that satisfies $\mathbf{T}^t = \mathbf{T}$.
A *anti-symmetric tensor* is a tensor that satisfies $\mathbf{T}^t = -\mathbf{T}$
Any tensor $\mathbf{T}$ can trivially be decomposed into a symmetric and an antisymmetric part. To see this, define

$$\begin{aligned}\mathbf{S}(\mathbf{T}) &= \mathbf{T} + \mathbf{T}^t \\ \mathbf{A}(\mathbf{T}) &= \mathbf{T} - \mathbf{T}^t\end{aligned} \tag{3.20}$$

Then $\mathbf{S}^t = \mathbf{S}$ is symmetric and $\mathbf{A}^t = -\mathbf{A}$ is anti-symmetric and

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

Finally, an important tensor is the *identity* or *unit* tensor, $\mathbf{1}$. It may be defined as the operator which acting on any vector, yields the vector itself. Evidently $\mathbf{1}$ is one of the special cases for which for all tensors $\mathbf{A}$

$$\mathbf{1} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{1}$$

If m is any scalar the product $m\mathbf{1}$ is called a *constant tensor* and has the property

$$(m\mathbf{1}) \cdot \mathbf{A} = \mathbf{A} \cdot (m\mathbf{1}) = m\mathbf{A}$$

Constant tensors therefore commute will all tensors and no other tensors have this property.

3.4 Tensor components in orthonormal bases

Just as is the case with vectors we can represent a tensor in terms of its components once we choose a basis. Let therefore $\mathcal{E}$ be an orthonormal basis and $\mathbf{T}$ be any tensor. Define

$$T_{ij} = \mathbf{e_i} \cdot \mathbf{T} \cdot \mathbf{e_j} \tag{3.21}$$

In words T_{ij} is the i'th component of the image of the j'th basis vector. T_{ij} is also called the ijth *component* of $\mathbf{T}$.

Using the linearity of $\mathbf{T}$ and scalar products one observes that the tensor is fully specified by these $3 \times 3 = 9$ quantities. Indeed, for any vector $\mathbf{c} = \sum_j c_j \mathbf{e_j}$ one obtains the i'th component of the image $\mathbf{u} = \mathbf{T} \cdot \mathbf{c}$ by

$$\begin{array}{ll} \mathbf{e_i} \cdot \left(\mathbf{T} \cdot (\sum_j c_j \mathbf{e_j}) \right) & = \\ \mathbf{e_i} \cdot \left(\sum_j c_j \mathbf{T} \cdot \mathbf{e_j} \right) & = \\ \sum_j c_j (\mathbf{e_i} \cdot \mathbf{T} \cdot \mathbf{e_j}) & = \sum_j T_{ij} c_j \end{array} \tag{3.22}$$

Thus,

$$\mathbf{T} \cdot \mathbf{c} = \sum_{ij} T_{ij} c_j \mathbf{e_i} \tag{3.23}$$

Since two indices are required for a full representation of $\mathbf{T}$, it is classified as a tensor of second rank. We will use the notation $[\mathbf{T}]_{\mathcal{E},ij}$ for the ij component of the tensor in basis $\mathcal{E}$ or simply $[\mathbf{T}]_{ij}$ when it is clear from the context (or irrelevant) which basis is implied. Note that the kl component of a dyad $\mathbf{e_i e_j}$ is

$$(\mathbf{e_i e_j})_{kl} = (\mathbf{e_k} \cdot \mathbf{e_i})(\mathbf{e_j} \cdot \mathbf{e_l}) = \delta_{ik} \delta_{jl}$$

and that the 3×3 dyad combinations $\mathbf{e_i e_j}$ form a basis for the tensor space

$$\mathbf{T} = \sum_{ij} T_{ij} \mathbf{e_i e_j} \tag{3.24}$$

Thus, the concepts of a dyadic and linear vector operator or tensor are identical and are equivalent to the concept of linear vector function in the sense that every linear vector function defines a certain tensor or dyadic, and conversely.

Conveniently, the components, T_{ij}, can be arranged in a *matrix* $\underline{\underline{T}} = (T_{ij})$,[1]

$$\underline{\underline{T}} = (T_{ij}) =_{\text{def.}} \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \tag{3.25}$$

In analogy with the notation used for vectors we write $\underline{\underline{T}} = [\mathbf{T}]_{\mathcal{E}}$ –or simply $\underline{\underline{T}} = [\mathbf{T}]$ when the basis is irrelevant– as a short hand notation for the component matrix of a tensor with respect to a given basis. Also, for the inverse operation (ie. the tensor obtained by expanding the components along the basis dyads) we shall use the notation

$$\mathbf{T} = (\underline{\underline{T}})_{\mathcal{E}} =_{\text{def}} \sum_{ij} T_{ij} \mathbf{e_i} \mathbf{e_j}. \tag{3.26}$$

The row-column convention in Eq. (3.25) implies that the representation of the dyad $\mathbf{e_i}\mathbf{e_j}$ is a matrix with a 1 in the i'th row and j'th column and zero everywhere else.

As for vectors one should notice the difference between a tensor and its matrix representation. The concept of a matrix is purely algebraic; matrices are arrays of numbers which may be added and multiplied according to particular rules (see below). The concept of a tensor is geometrical; a tensor may be represented in any particular coordinate system by a matrix, but the matrix must be transformed according to a definite rule if the coordinate system is changed.

3.4.1 Matrix algebra

Our definitions of the tensor-vector dot product and tensor-tensor dot-product are consistent with the ordinary rules of matrix algebra.

Tensor · vector

For instance, Eq. (3.22) shows that if $\mathbf{u} = \mathbf{T} \cdot \mathbf{c}$ then the relationship between the components of $\mathbf{u}$, $\mathbf{T}$ and $\mathbf{c}$ in any ortonormal basis will satisfy

$$u_i = [\mathbf{T} \cdot \mathbf{c}]_i = \sum_j T_{ij} c_j, \tag{3.27}$$

In standard matrix notation this is precisely equivalent to

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

[1] In accordance with standard notation in litterature the bracket around the components, (T_{ij}), is used to indicate the matrix collection of these. This is in analogy to the notation used for triplets $\underline{a} = (a_i)$, cf. section 2.3.

or

$$\underline{u} = \underline{T} \cdot \underline{a},$$

where the triplets, $\underline{u}$ and $\underline{a}$ are to be considered as a columns. Therefore, we have the correspondence between the tensor operation and its matrix representation

$$[\mathbf{T} \cdot \mathbf{c}] = [\mathbf{T}] \cdot [\mathbf{c}]$$

Tensor · tensor

Also, let us consider the implication of definition of the dot product of two tensors, Eq. (3.8) in terms of the components.

$$\begin{aligned} \mathbf{T} \cdot (\mathbf{S} \cdot \mathbf{c}) &= \mathbf{T} \cdot \left(\textstyle\sum_{kj} S_{kj} c_j \mathbf{e_k} \right) \\ &= \textstyle\sum_{kj} S_{kj} c_j \mathbf{T} \cdot \mathbf{e_k} \\ &= \textstyle\sum_{kj} S_{kj} c_j \sum_i T_{ik} \mathbf{e_i} \\ &= \textstyle\sum_{ij} \sum_k T_{ik} S_{kj} c_j \mathbf{e_i} \end{aligned} \tag{3.28}$$

Comparing this with Eq. (3.23), shows that

$$[\mathbf{T} \cdot \mathbf{S}]_{ij} = \sum_k T_{ik} S_{kj} \tag{3.29}$$

which again is identical to the ij'th element of the matrix product $\underline{T} \cdot \underline{S}$. A more transparent derivation is obtained by noting that for any vector $\mathbf{c}$

$$(\mathbf{e_k e_l}) \cdot \mathbf{c} = \mathbf{e_k} (\mathbf{e_l} \cdot \mathbf{c}) = \mathbf{e_k} c_l$$

and consequently

$$\begin{aligned} (\mathbf{e_i e_j} \cdot \mathbf{e_k e_l}) \cdot \mathbf{c} &= (\mathbf{e_i e_j}) \cdot (\mathbf{e_k e_l} \cdot \mathbf{c}) = (\mathbf{e_i e_j}) \cdot \mathbf{e_k} c_l = \mathbf{e_i} (\mathbf{e_j} \cdot \mathbf{e_k}) c_l \\ &= \mathbf{e_i} \delta_{jk} c_l = \delta_{jk} \mathbf{e_i} (\mathbf{e_l} \cdot \mathbf{c}) = \delta_{jk} (\mathbf{e_i e_l}) \cdot \mathbf{c} \end{aligned}$$

Therefore, one obtains the simple result for the dot product of two basis dyads

$$\mathbf{e_i e_j} \cdot \mathbf{e_k e_l} = \delta_{jk} \mathbf{e_i e_l} \tag{3.30}$$

Now, evaluating $\mathbf{T} \cdot \mathbf{S}$ in terms of its dyadic representations and collecting terms leads to

$$\begin{aligned} \mathbf{T} \cdot \mathbf{S} &= \left(\textstyle\sum_{ij} T_{ij} \mathbf{e_i e_j} \right) \cdot \left(\textstyle\sum_{kl} S_{kl} \mathbf{e_k e_l} \right) \\ &= \textstyle\sum_{ijkl} T_{ij} S_{kl} (\mathbf{e_i e_j}) \cdot (\mathbf{e_k e_l}) \\ &= \textstyle\sum_{ijkl} T_{ij} S_{kl} \delta_{jk} \mathbf{e_i e_l} \\ &= \textstyle\sum_{ij} \sum_l T_{il} S_{lj} \mathbf{e_i e_j} \end{aligned} \tag{3.31}$$

In obtaining the last expression we have exchanged the summation indices j and l. The final expression confirms Eq. (3.29). In summary,

$$[\mathbf{T} \cdot \mathbf{S}] = [\mathbf{T}] \cdot [\mathbf{S}]$$

Tensor + tensor

One can similarly show that the definition of the sum of two tensors Eq. (3.6),

$$(\mathbf{S}+\mathbf{T})\cdot\mathbf{c}=\mathbf{S}\cdot\mathbf{c}+\mathbf{T}\cdot\mathbf{c}, \quad (3.6)$$

implies that tensors are added by adding their component matrices. Indeed,

$$\begin{array}{ll}(\mathbf{S}+\mathbf{T})\cdot\mathbf{c} & =\left(\sum_{ij}S_{ij}\mathbf{e_i}\mathbf{e_j}+\sum_{ij}T_{ij}\mathbf{e_i}\mathbf{e_j}\right)\cdot\mathbf{c}\\ & =\sum_{ij}S_{ij}\mathbf{e_i}(\mathbf{e_j}\cdot\mathbf{c})+\sum_{ij}T_{ij}\mathbf{e_i}(\mathbf{e_j}\cdot\mathbf{c}) \quad \text{(def. Eq. (3.6))}\\ & =\sum_{ij}(S_{ij}+T_{ij})\mathbf{e_i}(\mathbf{e_j}\cdot\mathbf{c})\\ & =\left(\sum_{ij}(S_{ij}+T_{ij})\mathbf{e_i}\mathbf{e_j}\right)\cdot\mathbf{c}\end{array} \quad (3.32)$$

Consequently,

$$\mathbf{S}+\mathbf{T}=\sum_{ij}(S_{ij}+T_{ij})\mathbf{e_i}\mathbf{e_j} \quad (3.33)$$

or equivalently,

$$[\mathbf{S}+\mathbf{T}]_{ij}=S_{ij}+T_{ij}, \quad (3.34)$$

which means that the component matrix of $\mathbf{S}+\mathbf{T}$ is obtained by adding the component matrices of $\mathbf{S}$ and $\mathbf{T}$ respectively

$$[\mathbf{S}+\mathbf{T}]=[\mathbf{S}]+[\mathbf{T}]$$

It is left as an exercise to demonstrate that

$$[m\mathbf{T}]_{ij}=mT_{ij}, \quad (3.35)$$

where m is a scalar.

Transposition

Finally, the definition of the transpose of a tensor, Eq. (3.13), is also consistent with the algebraic definition of transposition. Specifically, comparing

$$\mathbf{c}\cdot\left(\sum_{ij}T_{ij}\mathbf{e_i}\mathbf{e_j}\right)=\sum_{ij}T_{ij}(\mathbf{c}\cdot\mathbf{e_i})\mathbf{e_j}=\sum_{ij}T_{ij}\mathbf{e_j}(\mathbf{e_i}\cdot\mathbf{c})=\left(\sum_{ij}T_{ji}\mathbf{e_i}\mathbf{e_j}\right)\cdot\mathbf{c} \quad (3.36)$$

with the definition of the transpose

$$\mathbf{T}^t\cdot\mathbf{c}=\mathbf{c}\cdot\mathbf{T}$$

the components satisfy

$$\left[\mathbf{T}^t\right]_{ij}=T_{ji}, \quad (3.37)$$

in agreement with the matrix definition of transposition. In other words, the matrix representation of the transposed tensor equals the transposed matrix representation of the tensor itself:

$$[\mathbf{T^t}]=[\mathbf{T}]^\mathbf{t}$$

Eq. (3.36) also demonstrates

$$[\mathbf{c}\cdot\mathbf{T}]_i = \sum_j c_j T_{ji} \tag{3.38}$$

In keeping with the convention that $\underline{c}$ represents a column, the matrix form of this equation is

$$\underline{c}^t \cdot \underline{T},$$

where $\underline{c}^t$ and and the image $\underline{c}^t \cdot \underline{T}$ will be rows.

3.4.2 Two-point components

As a curiosity, we mention that it is not compulsery to use the same basis for the left and the right basis-vectors when expressing a tensor in terms of its components. There are indeed cases where it is advantageous to use different left and right bases. Such tensor-components are called *two-point* components. In these cases one should make it clear in the notation that the first index of the component refer to a different basis than the second. For a two-point component matrix $\underline{\tilde{T}}$ we write

$$\mathbf{T} = (\underline{\tilde{T}})_{\mathcal{E}',\mathcal{E}} = \sum_{ij} \tilde{T}_{ij}\mathbf{e}'_\mathbf{i}\mathbf{e}_\mathbf{j}$$

and similarly, $[\mathbf{T}]_{\mathcal{E}'\mathcal{E},ij}$ for $\tilde{T}_{ij}$. Two-point components do not involve any extra formalism though. For instance, getting the components in the dyad basis $\mathcal{E}'\mathcal{E}'$ or in the basis $\mathcal{E}\mathcal{E}$ is only a question of respectively multiplying with the basis-vectors $\mathbf{e}'_\mathbf{i}$ to the right or the basis-vectors $\mathbf{e}_\mathbf{i}$ to the left on the expression $\sum_{kl}\tilde{T}_{kl}\mathbf{e}'_\mathbf{k}\mathbf{e}_\mathbf{l}$ and collect the terms. For instance

$$\begin{aligned}
[\mathbf{T}]_{\mathcal{E}\mathcal{E},ij} &= \mathbf{e}_\mathbf{i}\cdot\left(\textstyle\sum_{kl}\tilde{T}_{kl}\mathbf{e}'_\mathbf{k}\mathbf{e}_\mathbf{l}\right)\cdot\mathbf{e}_\mathbf{j} \\
&= \textstyle\sum_{kl}\tilde{T}_{kl}(\mathbf{e}_\mathbf{i}\cdot\mathbf{e}'_\mathbf{k})(\mathbf{e}_\mathbf{l}\cdot\mathbf{e}_\mathbf{j}) \\
&= \textstyle\sum_{kl}\tilde{T}_{kl}(\mathbf{e}_\mathbf{i}\cdot\mathbf{e}'_\mathbf{k})\delta_{lj} \\
&= \textstyle\sum_{k}\tilde{T}_{kj}(\mathbf{e}_\mathbf{i}\cdot\mathbf{e}'_\mathbf{k})
\end{aligned} \tag{3.39}$$

showing that that ij'th component in basis $\mathcal{E}\mathcal{E}$, T_{ij}, is

$$T_{ij} = \sum_k \tilde{T}_{kj}(\mathbf{e}_\mathbf{i}\cdot\mathbf{e}'_\mathbf{k})$$

3.5 Tensor fields

As for scalar and vector fields one speaks of a *tensor field* (here of rank 2) if to each position $\underline{x} = (x_1, x_2, x_3)$ of a region R in space there corresponds a tensor $\mathbf{T}(\underline{x})$. The best known examples of tensor fields in physics is the *stress* and *strain* tensor fields.

In cartesian coordinates a tensor field is given by

$$\mathbf{T}(\underline{x}) = \sum_{ij} T_{ij}(\underline{x}) \mathbf{e}_i \mathbf{e}_j \tag{3.40}$$

Note in particular that for each $i = 1, 2, 3$

$$\begin{array}{lll} \mathbf{a}_i(\underline{x}) & = \mathbf{T}(\underline{x}) \cdot \mathbf{e}_i & = \sum_j T_{ji}(\underline{x}) \mathbf{e}_j \quad \text{and} \\ \mathbf{b}_i(\underline{x}) & = \mathbf{e}_i \cdot \mathbf{T}(\underline{x}) & = \sum_j T_{ij}(\underline{x}) \mathbf{e}_j \end{array} \tag{3.41}$$

are *vector fields* with components $[\mathbf{a}_i]_j = T_{ji}$ and $[\mathbf{b}_i]_j = T_{ij}$, respectively.

3.5.1 Gradient, divergence and curl

Let a cartesian coordinate system be given $\mathcal{C} = (O, \mathcal{R})$.

Nabla operator

For any vector field, $\mathbf{a}(\mathbf{r})$ we can construct a tensor field by the nabla operator. Inserting $\mathbf{a}(\mathbf{r}) = \sum_i a_i(\mathbf{r}) \mathbf{e}_i$ "blindly" into the definition Eq. (2.41) gives

$$(\nabla \mathbf{a})(\mathbf{r}) = \sum_i \mathbf{e}_i \partial_i \left(\sum_j a_j(\mathbf{r}) \mathbf{e}_j \right) = \sum_{ij} (\partial_i a_j)(\mathbf{r}) \mathbf{e}_i \mathbf{e}_j \tag{3.42}$$

Assuming that ∇ is a proper vector operator[2] then $(\nabla \mathbf{a})(\mathbf{r})$ represents a tensor field with the components

$$[\nabla \mathbf{a}]_{ij}(\mathbf{r}) = (\partial_i a_j)(\mathbf{r}) \tag{3.43}$$

Notice that the components of $\nabla \mathbf{a}$ are that of an outer product between two ordinary vectors, $[\mathbf{ba}]_{ji} = b_j a_i$.

The tensor, $\nabla \mathbf{a}$, represents how the vector field $\mathbf{a}$ changes for a given displacement $\Delta \mathbf{r}$

$$\mathbf{a}(\mathbf{r} + \Delta \mathbf{r}) \approx \mathbf{a}(\mathbf{r}) + \Delta \mathbf{r} \cdot (\nabla \mathbf{a})(\mathbf{r})$$

or

$$d\mathbf{a} = d\mathbf{r} \cdot (\nabla \mathbf{a})$$

An alternative expression for the differential $d\mathbf{a}$ was obtained in Eq. (2.44),

$$d\mathbf{a} = d\mathbf{r} \cdot (\nabla \mathbf{a}) = (d\mathbf{r} \cdot \nabla) \mathbf{a}$$

Notice that this identity is in accordance with the usual definition for a dyad operating on a vector from the left.

[2] We recall that the ∇ operator has been defined relative to a chosen coordinate system. It remains to be proven that operators derived from ∇ such as gradients, curls, divergence etc. "behave as" proper scalars or vectors. We return to this point in 4.5.

Divergence

The divergence of a tensor field is obtained straight forwardly by applying the definition of ∇ and the *dot*-product. With respect to a cartesian basis $\mathcal{R}$ we have

$$\begin{array}{ll} \big(\nabla \cdot \mathbf{T}\big)(\underline{x}) & = \left(\sum_i \mathbf{e}_i \partial_i\right) \cdot \left(\sum_{jk} T_{jk}(\underline{x}) \mathbf{e}_j \mathbf{e}_k\right) \\ & = \sum_i \sum_k \partial_i T_{ik}(\underline{x}) \mathbf{e}_k \\ & = \sum_k \left(\sum_i \partial_i T_{ik}(\underline{x})\right) \mathbf{e}_k. \end{array} \tag{3.44}$$

Consequently, $\mathbf{a} = \nabla \cdot \mathbf{T}$ is a *vector* with components $a_k = \sum_i \partial_i T_{ik}$, jvf. Eq. (3.41). The result is actually easier seen in pure subscript notation

$$[\nabla \cdot \mathbf{T}]_k = \sum_i \partial_i T_{ik},$$

because it follows directly from the matrix algebra, Eq. (3.38). Another way of viewing the divergence of a tensor field is to take the divergence of each of the vector fields $\mathbf{a}_i = \mathbf{T} \cdot \mathbf{e}_i$

$$\nabla \cdot \mathbf{T} = \sum_i \left(\nabla \cdot \mathbf{a}_i\right) \mathbf{e}_i$$

Curl

The curl of a tensor field is obtained similarly to the divergence. The result is most easily obtained by applying the algebraic definition of curl, Eq. (2.47). Then we obtain

$$[\nabla \times \mathbf{T}]_{ij} = \sum_{mn} \epsilon_{imn} \partial_m T_{nj}$$

Thus the curl of a tensor field is another tensor field. As for the divergence we obtain the same result by considering the curl operator on each of the three vector field $\mathbf{a}_j = \mathbf{T} \cdot \mathbf{e}_j$

$$\nabla \times \mathbf{T} = \sum_j \big(\nabla \times \mathbf{a}_j\big) \mathbf{e}_j,$$

where an outer vector product is involved in each of the terms $\big(\nabla \times \mathbf{a}_j\big)\mathbf{e}_j$.

3.5.2 Integral theorems

The two important theorems for vector fields, Gauss theorem and Stokes theorem, can be directly translated to tensor fields. The analogy is most easily seen in tensor notation. Then for a vector field $\mathbf{a}(\mathbf{r})$ the two theorem reads

$$\begin{array}{lll} \int_V \left(\sum_i \partial_i a_i\right) dV & = \oint_S \left(\sum_i a_i n_i\right) dS & \text{Gauss} \\ \int_V \left(\sum_{ijk} \epsilon_{ijk} \partial_j a_k\right) n_i dS & = \oint_C \left(\sum_i a_i dx_i\right) & \text{Stokes} \end{array} \tag{3.45}$$

The corresponding tensor versions are then

$$\begin{array}{lll} \int_V \left(\sum_i \partial_i T_{il}\right) dV & = \oint_S \left(\sum_i T_{il} n_i\right) dS & \text{Gauss} \\ \int_V \left(\sum_{ijk} \epsilon_{ijk} \partial_j T_{kl}\right) n_i dS & = \oint_C \left(\sum_i T_{il} dx_i\right) & \text{Stokes} \end{array} \tag{3.46}$$

Here l refer to any component index $l = 1, 2, 3$. The reason these formulas also works for tensors is that for a fixed l, $\mathbf{a}_l = \mathbf{T} \cdot \mathbf{e}_l = \sum_i T_{il} \mathbf{e}_i$ defines a vector field, c.f. (3.41), and Gauss' and Stokes' theorems works for each of these.

Due to Gauss theorem the physical interpretation of the divergence of a tensor field is analogous to the divergence of a vector field. For instance, if a flow field $\mathbf{v}(\mathbf{r}, t)$ is advecting a vector $\mathbf{a}(\mathbf{r}, t)$ then the outer product $\mathbf{J}(\mathbf{r}, t) = \mathbf{v}(\mathbf{r}, t)\mathbf{a}(\mathbf{r}, t)$ is a tensor field where $[\mathbf{J}]_{ij}(\mathbf{r}, t) = v_i(\mathbf{r}, t) a_j(\mathbf{r}, t)$ describes how much of the j'th component of the vector $\mathbf{a}$ is transported in direction $\mathbf{e}_i$. The divergence $\nabla \cdot \mathbf{J}$ is then a vector where each component $[\nabla \cdot \mathbf{J}]_j$ corresponds to the accumulation of a_j in an infinitesimal volume due to the flow. In other words, if $\mathbf{a}$ represents a conserved quantity then we have a *continuity equation* for the vector field $\mathbf{a}$

$$\frac{\partial \mathbf{a}}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

Chapter 4

Tensor calculus

In the two preceeding chapters we have developed the basic algebra to for scalar, vector and rank 2 tensors. We have seen that the notation of vectors and tensors comes in two flavours. The first notation insists in not refering to coordinates or components at all. This geometrically defined notation, called *direct notation*, is explictly invariant to the choise of coordinate system and is the one adopted in the first part of chapter two and three (2.1-2.2, 3.1-3.3). The second notation, based on components and indices, is called *tensor notation* and is the one that naturally arises with a given coordinate system. Here, vectorial or tensorial relations are expressed algebraically.

Seemingly, a discrepancy exists between the two notations in that the latter appears to depend on the choise of coordinates. In section 4.2 we remove this descrepancy by learning how to transform the components of vectors and tensors when the coordinate system is changed. As we shall see these transformation rules guarantee that one can unambigously express vectorial or tensorial relations using component/tensor notation because the expressions will preserve the form upon a coordinate transformation. Indeed, we should already expect this to be the case since we have not made any specific assumptions about the coordinate system in the propositions regarding relations between vector or tensor components already presented, except of this system beeing rectangular. It is possible to generalize the tensor notation to ensure that the formalism takes the same form in non-rectangular coordinate systems as well. This requirement is known as *general covariance*. In the following we will, however, restrict the discussion to rectangular coordinates.

Since tensor notation often requires knowledge of transformation rules between coordinate systems one may validly ask why use it at all. First of all, in quantifying any physical system we will eventually have to give a numerical representation which for vectors or tensors imply to specify the value of their components with respect to some basis. Secondly, in many physical problems it is natural to introduce tensors of rank higher than two[1]. To insist on a direct

[1]The Levi-Civita symbol being an example of a (pseudo)-tensor of rank 3.

notation for these objects become increasingly tedious and unnatural. In tensor notation no new notation needs to be defined, only those we have already seen (some in disguise), which is reviewed in section 4.3. It is therefore strongly recommended to become familiar with the tensor notation once and for all.

4.1 Tensor notation and Einsteins summation rule

Using tensor notation an index will either be *free* or *bound*. A free index occurs exactly once in a term and has the meaning "for any component". For instance, in the formula for the addition of two vectors in terms of its components, Eq. (2.13),

$$[\mathbf{a}+\mathbf{b}]_i = [\mathbf{a}]_i + [\mathbf{b}]_i,$$

index i is free. The same free indices must occur in each term. As in the above example, an equation with exactly one free index in each term expresses a vector identity, an expression with two free indices in each term expresses a tensor identity, etc.

A bound index (or a dummy index) refers to a summation index. For instance the scalar product in an orthonormal basis, Eq. (2.18), reads

$$\mathbf{a}\cdot\mathbf{b} = \sum_i a_i b_i.$$

Here, i is a bound index and can be renamed without changing the meaning of the expression. The situation where a dummy index appears exactly twice in a product occurs so often that it is convenient to introduce the convention that the summation symbol may be left out in this specific case. In particular, it occurs in any component representation of a "dot" or inner product. Hence we may write

$$\begin{array}{llll}
\mathbf{a}\cdot\mathbf{b} & = \sum_i a_i b_i & = a_i b_i & (i \text{ a bound index}) \\
[\mathbf{T}\cdot\mathbf{a}]_j & = \sum_i T_{ji} a_i & = T_{ji} a_i & (i \text{ a bound index, } j \text{ a free index}) \\
[\mathbf{T}\cdot\mathbf{S}]_{ij} & = \sum_k T_{ik} S_{kj} & = T_{ik} S_{kj} & (ij \text{ free indices, } k \text{ bound index}) \\
\mathbf{T}:\mathbf{S} & = \sum_{ij} T_{ij} S_{ij} & = T_{ij} S_{ij} & ij \text{ bound indices}
\end{array} \tag{4.1}$$

Similar, for the expansion of the components along the basis vectors, Eq. (2.10)

$$\mathbf{a} = \sum_i a_i \mathbf{g_i} = \mathbf{a_i g_i}$$

and for the divergence of a vector field

$$\nabla\cdot\mathbf{a} = \sum_i \partial_i a_i = \partial_i a_i$$

The convention of implicit summation which is very convenient for tensor calculus is due to Einstein and is referred to as Einsteins summation rule. *In this and the following chapters we shall adopt the implicit summation rule unless otherwise stated.*

4.2 Orthonormal basis transformation

Let $\mathcal{E}$ and $\mathcal{E}'$ be two different orthonormal bases. The basis vectors of the primed basis may be revolved into the unprimed basis

$$\mathbf{e}'_j = a_{ji}\mathbf{e}_i \qquad \text{(sum over } i\text{)}, \tag{4.2}$$

where

$$a_{ji} =_{\text{def}} \mathbf{e}'_j \cdot \mathbf{e}_i \tag{4.3}$$

represents the cosine of the angle between $\mathbf{e}'_j$ and $\mathbf{e}_i$. The relation between the primed and unprimed components of any (proper) vector

$$\mathbf{v} = c_i\mathbf{e}_i = c'_j\mathbf{e}'_j$$

can be obtained by dotting with $\mathbf{e}'_j$ or $\mathbf{e}_i$

$$v'_j = a_{ji}v_i \tag{4.4}$$
$$v_i = a_{ji}v'_j \tag{4.5}$$

Here, $v'_j = [\mathbf{v}]_{\mathcal{E}',j}$ and $v_i = [\mathbf{v}]_{\mathcal{E},i}$. Note, that the two transformations differ in whether the sum is over the first or the second index of a_{ij}. The matrix version of Eq. (4.4) reads

$$\underline{v}' = \underline{A} \cdot \underline{v}, \qquad \underline{A} = (a_{ij}), \tag{4.6}$$

and Eq. (4.5)

$$\underline{v} = \underline{A}^t \cdot \underline{v}'.$$

Using the same procedure the primed and unprimed components of a tensor

$$\mathbf{T} = T'_{ij}\mathbf{e}'_\mathbf{i}\mathbf{e}'_\mathbf{k} = T_{jl}\mathbf{e}_\mathbf{j}\mathbf{e}_\mathbf{l}$$

[2] are related by

$$T'_{ik} = \mathbf{e}'_i \cdot \mathbf{T} \cdot \mathbf{e}'_k = a_{ij}a_{kl}T_{jl} \tag{4.7}$$
$$T_{ik} = \mathbf{e}_i \cdot \mathbf{T} \cdot \mathbf{e}_k = a_{ij}a_{kl}T_{jl}, \tag{4.8}$$

where $T'_{ik} = [\mathbf{T}]_{\mathcal{E}',ik}$ and $T_{ik} = [\mathbf{T}]_{\mathcal{E},ik}$. Note that the transformation matrix $\underline{A}$ is simply the two-point components of the identity tensor

$$\underline{A} = [\mathbf{1}]_{\mathcal{E}'\mathcal{E}}$$

[2] In keeping with Einsteins summation rule this expression is short hand notation for

$$\mathbf{T} = \sum_{ik} T'_{ij}\mathbf{e}'_\mathbf{i}\mathbf{e}'_\mathbf{k} = \sum_{jl} T_{jl}\mathbf{e}_\mathbf{j}\mathbf{e}_\mathbf{l}$$

4.2.1 Cartesian coordinate transformation

A particular case of an orthonormal basis transformation is the transformation between two cartesian coordinate systems, $\mathcal{C} = (O, \mathcal{R})$ and $\mathcal{C}' = (O', \mathcal{R}')$. Here, the form of the vector transformation, Eq. (4.4), can be directly translated to the coordinates themselves. To show this we employ the simple relation between coordinates and the position vector, Eq. (2.9), valid in any cartesian system. The displacement vector between any two points is given by $\Delta\mathbf{r} = \Delta x_i \mathbf{e}_i$. This vector –or its differential analogue $d\mathbf{r} = dx_i \mathbf{e}_i$ – is a proper vector independent of the coordinate system. Following relation between the coordinate differentials dx_i and dx'_j must then be satisfied

$$d\mathbf{r} = dx_i \mathbf{e}_i = dx'_j \mathbf{e}'_j$$

Multiplying with $\mathbf{e}'_\mathbf{j}$ we obtain

$$dx'_j = dx_i \mathbf{e}_\mathbf{i} \cdot \mathbf{e}'_\mathbf{j}$$

or equivalently

$$\frac{\partial x'_j}{\partial x_i} = \mathbf{e}'_\mathbf{j} \cdot \mathbf{e}_\mathbf{i} = a_{ji}, \tag{4.9}$$

where a_{ji} is defined as in Eq. (4.3). Since the rhs. is constant (constant basis vectors) the integration of Eq. (4.9) gives

$$x'_j = a_{ji} x_i + d_j. \tag{4.10}$$

Here, d_j is the j'th coordinate of the origin, O, of the unprimed coordinate system as seen from the primed coordinate system, $\underline{d} = (\underline{x}')_{\mathcal{C}'}(O)$. In matrix notation Eq. (4.10) takes the form

$$\underline{x}' = \underline{A} \cdot \underline{x} + \underline{d}, \qquad \underline{A} = (a_{ij}), \tag{4.11}$$

which is identical to Eq. (4.6) except from the optional displacement.

4.2.2 The orthogonal group

Considering Eq. (4.4) and Eq. (4.5) as a set of two matrix equations we note that the transpose matrix $\underline{A}^t$ operates as the inverse of $\underline{A}$

$$\underline{A} \cdot \underline{A}^t = \underline{1}, \tag{4.12}$$

where $\underline{1}$ is the identity matrix, $(\underline{1})_{ij} = \delta_{ij}$. In index notation

$$(\underline{A} \cdot \underline{A}^t)_{ij} = a_{ik} a_{kj} = \delta_{ij} \tag{4.13}$$

A matrix with the above property is called *orthogonal*, and the set of of all orthogonal 3×3 matrices constitutes a continuous group[3] called $O(3)$. According to Eq. (4.12)

$$\det(\underline{A} \cdot \underline{A}^t) = \det(\underline{A})^2 = 1,$$

[3]Recall, that a mathematical group $(G, *)$ is a set G with a binary operator, $*$, that satisfies following axioms:

so

$$\det(\underline{A}) = \pm 1$$

Transformations with the determinant $+1$ or -1 cannot be continuously connected. Therefore, the set of transformations with determinant $+1$ itself forms a group, called $SO(3)$, which represents the set of all rotations. By considering a pure *reflection*, $\underline{R} = (r_{ij})$, in the origin of a cartesian coordinate system

$$x_i' = -x_i = r_{ij}x_j, \qquad r_{ij} = -\delta_{ij}$$

we see that $\det(\underline{R}) = -1$. Clearly, $\underline{R} \cdot \underline{R} = \underline{1}$, so the set $Z(2) = \{1, R\}$ also forms a group. Consequently, we may write $O(3) = Z(2) \otimes SO(3)$. In words, any orthogonal matrix can be decomposed into a pure rotation and an optional reflection. Note, that any matrix with $\det = -1$ will change the handedness of the coordinate system.

4.2.3 Algebraic invariance

In view of the fact that the various vector and tensor operations (section 2.1-2.3, 3.1-3.3) were defined without reference to a coordinate system, it is clear that all algebraic rules for computing sums, products, transposes, etc. of vectors and tensors (cf. section 3.4.1) will be unaffected by an orthogonal basis transformation. Thus for examples

$$\begin{array}{ll} \mathbf{a} \cdot \mathbf{b} & = a_i' b_i' = a_j b_j \\ [\mathbf{a} + \mathbf{b}]_{\mathcal{E}',i} & = a_i' + b_i' \\ [\mathbf{T} \cdot \mathbf{c}]_{\mathcal{E}',i} & = T_{ij}' c_j' \\ [\mathbf{T}^t]_{\mathcal{E}',ij} & = T_{ji}', \end{array} \tag{4.14}$$

where, $a_i' = [\mathbf{a}]_{\mathcal{E}',i}$, $a_j = [\mathbf{a}]_{\mathcal{E},j}$, $T_{ji}' = [\mathbf{T}]_{\mathcal{E}',ji}$ etc.

Any property or relation between vectors and tensors which is expressed in the same algebraic form in all coordinate systems has a geometrical meaning independent of the coordinate system and is called an *invariant* property or relation.

The invariance of the algebraic rules for vector and tensor operations can be verified directly by using the orthogonality property of the transformation matrix, Eq. (4.13). In particular, whenever a "dot" or inner product appears between two vectors, a vector and a tensor or two tensors we can pull together two tranformation matrices and exploit Eq. (4.13). For instance, to verify the

1. *Closure*: $\forall a, b \in G : a * b \in G$.
2. *Associativity*: $\forall a, b, c \in G : (a * b) * c = a * (b * c)$
3. *Identity element*: There exists an element $e \in G$ so that $\forall a : a * e = e * a = a$.
4. *Inverse element*: For each $a \in G$ the exists an inverse $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

Orthogonal matrices forms a group with matrix multiplication as the binary operator and the identity matrix as the identity element. Notice in particular that the composition or multiplication of two orthogonal transformations is orthogonal.

third invariance above, ie. that the i'th component of the image $\mathbf{u} = \mathbf{T} \cdot \mathbf{c}$ of $\mathbf{T}$ operating on $\mathbf{c}$, is always obtained as

$$u_i = T_{ij} c_j, \qquad (3.27)$$

irrespective of the chosen coordinate system, we can tranform the components of $\mathbf{u}$ from one (unprimed) to another (primed) coordinate system and see if these equal the transformed component matrix of $\mathbf{T}$ times the transformed component triplet of $\mathbf{c}$:

$$\begin{aligned}
[\mathbf{T}]_{\mathcal{E}',ij}[\mathbf{c}]_{\mathcal{E}',j} &= T'_{ij} c'_j \\
&= (a_{ik} a_{jl} T_{kl})(a_{jm} c_m) && \text{Transf. rules for components} \\
&= a_{ik} T_{kl} (a_{jl} a_{jm}) c_m \\
&= a_{ik} T_{kl} \delta_{ml} c_m && \text{Orthogonality (summing over } k\text{)} \\
&= a_{ik} (T_{kl} c_l) \\
&= a_{ik} [\mathbf{u}]_{\mathcal{E},k} \\
&= [\mathbf{u}]_{\mathcal{E}',i} && \text{Tranf. rules for components}
\end{aligned} \qquad (4.15)$$

We shall return to other important invariances in chapter 5

4.2.4 Active and passive transformation

Comparing the transformation of the components of a vector upon a basis transformation, Eq. (4.4) with the matrix representation of a tensor operating on a vector, Eq. (3.27), it is clear that these two operations are defined in the same manner algebraically. Since the first equation only expresses the change of components due to a change of basis (the vector itself remains unchanged) it is referred to as a *passive* transformation as opposed to the second case which reflects an *active* transformation of the vector. Here, we shall show a simple algebraic relation between the two type of transformations. Let $\mathbf{O}$ represent an *orthogonal tensor*, ie. a tensor for which

$$\mathbf{O}^t \cdot \mathbf{O} = \mathbf{1},$$

where $\mathbf{1}$ as usual denotes the identity tensor. In any orthonormal basis the matrix representation of this identity will be that of Eq. (4.12). Further, let a new basis system $\mathcal{E}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ be defined by

$$\mathbf{e}'_i = \mathbf{O} \cdot \mathbf{e}_i$$

where $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the old basis. We shall use the short-hand notation

$$\mathcal{E}' = \mathbf{O}(\mathcal{E})$$

Due to the orthogonality property of $\mathbf{O}$, $\mathcal{E}'$ will be ortonormal iff $\mathcal{E}$ is.

The components of a vector $\mathbf{v}$ in the primed basis relates to the the unprimed components as

$$\begin{aligned} v'_i &= \mathbf{e}'_i \cdot \mathbf{v} \\ &= (\mathbf{O} \cdot \mathbf{e}_i) \cdot \mathbf{v} \\ &= [\mathbf{O} \cdot \mathbf{e}_i]_{\mathcal{E},j} [\mathbf{v}]_{\mathcal{E},j} \qquad \text{(Scalar product wrt. } \mathcal{E}) \\ &= [\mathbf{O}]_{jk} [\mathbf{e}_i]_k v_j \qquad \text{(Eq. (3.27), all components wrt. } \mathcal{E}) \\ &= [\mathbf{O}]_{jk} \delta_{ik} v_j \\ &= [\mathbf{O}]_{ji} v_j \\ &= [\mathbf{O}^t]_{ij} v_j \\ &= (\underline{O}^t \cdot \underline{v})_i \end{aligned} \tag{4.16}$$

This shows that the matrix, $\underline{A}$, representing the basis transformation, Eq. (4.6), satisfies

$$\underline{A} = [\mathbf{O}^t]_{\mathcal{E}} = \underline{O}^t = \underline{O}^{-1}.$$

Being explicit about the bases involved we may write this identity as

$$[\mathbf{1}]_{\mathcal{E}'\mathcal{E}} = [\mathbf{O}^{-1}]_{\mathcal{E}}, \qquad \text{where} \qquad \mathcal{E}' = \mathbf{O}(\mathcal{E}), \tag{4.17}$$

which is to say that the matrix representing a passive transformation from one to another basis, $\mathcal{E} \to \mathcal{E}'$, equals the matrix representation wrt. $\mathcal{E}$ of the tensor mapping the new basis vectors onto the old ones, $\mathbf{O}^{-1}(\mathcal{E}') = \mathcal{E}$.

4.2.5 Summary on scalars, vectors and tensors

To summarize the present section we may give an *algebraic definition* of scalars, vectors and tensors by refering to the transformation properties of their components upon a change of coordinates. For an cartesian coordinate transformation, Eq. (4.10) following transformation properties of scalars, vectors and tensors apply.

Scalars

A quantity m is a *scalar* if it has the same value in every coordinate system. Consequently it transforms according to the rule

$$m' = m$$

Vectors

A triplet of real numbers, $\underline{v}$, is a *vector* if the components transform according to

$$v'_j = a_{ji} v_i$$

The matrix notation of the transformation rule is

$$\underline{v}' = \underline{A} \cdot \underline{v}$$

Rank 2 Tensors

A rank two tensor is a 3×3 matrix of real numbers, $\underline{T}$, which transforms as the outer product of two vectors

$$T'_{ij} = a_{ik} a_{jl} T_{kl}$$

Again it should be noticed that it is the transformation rule which guarantees that the nine quantities collected in a matrix form a tensor. The matrix notation of this transformation is

$$\underline{T}' = \underline{A} \cdot \underline{T} \cdot \underline{A}^t$$

4.3 Tensors of any rank

4.3.1 Introduction

Tensors of rank higher than two are obtained by a straightforward iteration of the definition of the outer product. For instance the outer product between a dyad $\mathbf{a}_1\mathbf{a}_2$ and a vector $\mathbf{a}_3$ can be defined as the linear operator, $\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$, that maps any vector $\mathbf{c}$ into a tensor according to the rule

$$\mathbf{a_1}\mathbf{a_2}\mathbf{a_3}(\mathbf{c}) =_{\text{def.}} \mathbf{a_1}\mathbf{a_2}(\mathbf{a_3} \cdot \mathbf{c}) \tag{4.18}$$

In words, for each $\mathbf{c}$ the object $\mathbf{a_1}\mathbf{a_2}\mathbf{a_3}$ associates a tensor obtained by multiplying the dyad $\mathbf{a_1}\mathbf{a_2}$ with the scalar $\mathbf{a_3} \cdot \mathbf{c}$. Definition Eq. (4.18) preserves the invariance with respect to the choise of coordinates. Addition of triple products and scalar multiplication is defined analogous to the algebra for dyads.

In section 3.4 we demonstated that a basis for second order tensors is obtained by forming all 3×3 dyad combinations, $\mathbf{e}_i\mathbf{e}_j$, between orthonormal basis vectors. Similarly, one can demonstrate that a basis for all linear operators, $\mathbf{T}_3$, which map a vector into a second rank tensor is obtained by forming the 3^3 combinations of the direct products $\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k$. Consequently, for any $\mathbf{T}_3$ there exists a unique set of quantities, T_{ijk}, such that

$$\mathbf{T}_3 = T_{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k \qquad \text{NB! Einstein summation convention}$$

Not surprisingly, these quantities, T_{ijk}, are called the *components* of $\mathbf{T}_3$ and since three indicies are needed, $\mathbf{T}_3$ is called a third rank tensor. Upon a coordinate transformation we would need to transform the components of each individual vector in the triple product. Expressing the original basis vectors in terms of the new basis

$$\mathbf{e}_i = a_{ji}\mathbf{e}'_j$$

one obtains

$$\begin{aligned} \mathbf{T}_3 &= T'_{lmn}\mathbf{e}'_l\mathbf{e}'_m\mathbf{e}'_n = T_{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k \\ &= T_{ijk}(a_{li}\mathbf{e}'_l)(a_{mj}\mathbf{e}'_m)(a_{nk}\mathbf{e}'_n) = a_{li}a_{mj}a_{nk}T_{ijk}\mathbf{e}'_l\mathbf{e}'_m\mathbf{e}'_n \end{aligned} \tag{4.19}$$

This shows that upon an orthogonal transformation the components of a third rank tensor transform as

$$T'_{lmn} = a_{li}a_{mj}a_{nk}T_{ijk}$$

The inverse transformation follows from expressing the new basis vectors in terms of the old one, $\mathbf{e}'_j = a_{ji}\mathbf{e}_i$ in Eq. (4.19)

$$T_{ijk} = a_{li}a_{mj}a_{nk}T'_{lmn}.$$

We can continue this procedure, defining the outer product of four vectors in terms triple products, cf. Eq. (4.18), and introducing addition and scalar multiplication of these objects as well.

In keeping with previous notation, the set of all components is written as

$$\underline{T} = (T_{ijk}).$$

Here, ijk are dummy indices and the bracket around T_{ijk} indicates the set of all of these

$$(T_{ijk}) = \{\ T_{ijk} \mid i = 1,2,3\ ;\ j = 1,2,3\ ;\ k = 1,2,3\ \}.$$

Also, the tensor is obtained from the components as

$$\mathbf{T} = (\underline{T})\mathcal{E} = (\ (T_{ijk})\ )\mathcal{E} =_{\text{def}} T_{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k$$

4.3.2 Definition

In general, a tensor $\mathbf{T}$ of rank r is defined as a set of 3^r quantities, $T_{i_1 i_2 \cdots i_r}$, that upon an orthogonal change of coordinates, Eq. (4.11), transform as the outer product of r vectors

$$T'_{j_1 j_2 \cdots j_r} = a_{j_1 i_1}a_{j_2 i_2} \cdots a_{j_r i_r}T_{i_1 i_2 \cdots i_r}. \tag{4.20}$$

[4] Accordingly, a vector is a rank 1 tensor and a scalar is a rank 0 tensor.

It is common to refer to both $\mathbf{T}$ and the set $(T_{i_1 i_2 \cdots i_r})$ as a tensor, although the latter quantities are strictly speaking only the *components of the tensor* with respect to some particular basis $\mathcal{E}$. However, the distinction between a tensor and its components becomes unimportant, provided we keep in mind that $T'_{j_1 j_2 \cdots j_r}$ in Eq. (4.20) are components of the *same* tensor only with respect to a different basis $\mathcal{E}'$. It is precisely the transformation rule, Eq. (4.20), that ensures that the two set of components represent the same tensor.

For completeness, the inverse transformation is obtained by summing on the first indices of the transformation matrix elements instead of the second ones, cf. section 4.3.2:

$$T_{i_1 i_2 \cdots i_r} = a_{j_1 i_1}a_{j_2 i_2} \cdots a_{j_r i_r}T'_{j_1 j_2 \cdots j_r}.$$

[4] The notation $T_{i_1 i_2 \cdots i_r}$ may be confusing at first sight. However, $i_1, i_2, \cdots i_r$ are simply r independent indices each taking values in the range $\{1,2,3\}$. Choosing a second rank tensor as an example and comparing with previous notation, T_{ij}, it simply means that $i_1 = i$, $i_2 = j$. For any particular component of the tensor, say T_{31}, we would have with the old notation $i = 3, j = 1$ and in the new notation $i_1 = 3, i_2 = 1$.

4.3.3 Basic algebra

Tensors can be combined in various ways to form new tensors. Indeed, we have already seen several examples thereof[5]. The upshot of these rules below is that tensor operations that look "natural" are permissible in the sense that if one starts with tensors the result will be a tensor **??**.

Addition/substraction

Tensors of the same rank may be added together. Thus

$$C_{i_1 i_2 \cdots i_r} = A_{i_1 i_2 \cdots i_r} + B_{i_1 i_i 2 \cdots i_r}$$

is a tensor of rank r if $\mathbf{A}$ and $\mathbf{B}$ are tensors of rank r. This follows directly from the linearity of Eq. (4.20),

$$\begin{aligned} C'_{j_1 j_2 \cdots j_r} &= A'_{j_1 j_2 \cdots j_r} + B'_{j_1 j_2 \cdots j_r} \\ &= a_{j_1 i_1} a_{j_2 i_2} \cdots a_{j_r i_r} A_{i_1 i_2 \cdots i_r} + a_{j_1 i_1} a_{j_2 i_2} \cdots a_{j_r i_r} B_{i_1 i_2 \cdots i_r} \\ &= a_{j_1 i_1} a_{j_2 i_2} \cdots a_{j_r i_r} \left(A_{i_1 i_2 \cdots i_r} + B_{i_1 i_2 \cdots i_r} \right) \\ &= a_{j_1 i_1} a_{j_2 i_2} \cdots a_{j_r i_r} C_{i 1 i_2 \cdots i_r}. \end{aligned}$$

Consequently, the 3^r quantities $(C_{i_1 i_2 \cdots i_r})$ transform as the components of r rank tensor when both $(A_{i_1 i_2 \cdots i_r})$ and $(B_{i_1 i_i 2 \cdots i_r})$ do (We made use of this in the second line of the demonstation above). It follows directly that substraction of two equally ranked tensors is also a tensor of the same rank, and that tensorial addition/substraction is commutative and associative.

Outer product

The *product* of two tensors is a tensor whose rank is the sum of the ranks of the given tensors. This product which involves ordinary multiplication of the components of the tensor is called the *outer product*. It is the natural generalization of the outer product of two vectors (two tensors of rank 1) defined in section 3.2. For example

$$C_{i_1 i_2 \cdots i_r j_1 j_2 \cdots j_s} = A_{i_1 i_2 \cdots i_r} B_{j_1 j_2 \cdots j_s} \tag{4.21}$$

is a tensor of rank $r+s$ if $A_{i_1 i_2 \cdots i_r}$ is a tensor of rank r and $B_{j_1 j_2 \cdots j_s}$ is a tensor of rank s. Note, that this rule is consistent with the cartesian components of a dyad $\mathbf{c} = \mathbf{ab}$ in the specific case where $\mathbf{A} = \mathbf{a}$ and $\mathbf{B} = \mathbf{b}$ are vectors. Also, a tensor may be multiplied with a scalar $m = B$ (tensor of rank 0) according to the same rule

$$C_{i_1 i_2 \cdots i_r} = m A_{i_1 i_2 \cdots i_r}$$

Note, that not every tensor can be written as a product of two tensors of lower rank. We have already emphasized this point in case of second rank tensors which can not in general be expressed as a single dyad. For this reason divison of tensors is not always possible.

[5] The composition, $\mathbf{T} \cdot \mathbf{S}$, of two second rank tensors, $\mathbf{T}$ and $\mathbf{S}$, forms a new second rank tensor. The operation of a second rank tensor on a vector (1. rank tensor) gives a new vector. The addition of two second rank tensors gives a new second rank tensor, etc.

Permutation

It is permitted to exchange or permute indices in a tensor and still remain with a tensor. For instance, if we define a new set of 3^r quantities $C_{i_1 i_2 \cdots i_r}$ from a tensor $A_{i_1 i_2 \cdots i_r}$ by permuting two arbitrarily chosen index numbers α and $\beta > \alpha$:

$$C_{i_1 i_2 \cdots i_{\alpha-1} i_\alpha i_{\alpha+1} \cdots i_{\beta-1} i_\beta i_{\beta+1} \cdots i_r} = A_{i_1 i_2 \cdots i_{\alpha-1} i_\beta i_{\alpha+1} \cdots i_{\beta-1} i_\alpha i_{\beta+1} \cdots i_r},$$

this new set will also be a tensor. Its tensorial property follows from the symmetry among the a-factors in Eq. (4.20). Tensors obtained from permutting indices are called *isomers*. Tensors of rank less than two (ie. scalars and vectors) have no isomers. A tensor of rank two has precisely one isomer, the transposed one, obtained by setting $\alpha = 1$ and $\beta = 2$ in the above notation:

$$(\mathbf{T}^t)_{i_1 i_2} = T_{i_2 i_1}$$

Contraction

The most important rule in tensor algebra is the *contraction rule* which states that if two indices in a tensor of rank $r + 2$ are put equal and summed over, then the result is again a tensor with rank r. Because of the permutation rule discussed above we only have to demonstrate it for the first two indices. The contraction rule states that if $\mathbf{A}$ is a tensor of rank $r + 2$ then

$$B_{j_1 j_2 \cdots j_r} = A_{iij_1 j_2 \cdots j_r} \qquad (\text{NB!} \quad A_{iij_1 j_2 \cdots j_r} = \textstyle\sum_{i=1}^{3} A_{iij_1 j_2 \cdots j_r} \ )$$

is also a tensor of rank r. The proof follows

$$\begin{aligned} B'_{j_1 j_2 \cdots j_r} &= A'_{iij_1 j_2 \cdots j_r} \\ &= a_{ik} a_{il} a_{j_1 m_1} a_{j_2 m_2} \cdots a_{j_r m_r} A_{klm_1 m_2 \cdots m_r} \\ &= \delta_{kl} a_{j_1 m_1} a_{j_2 m_2} \cdots a_{j_r m_r} A_{klm_1 m_2 \cdots m_r} A_{kkm_1 m_2 \cdots m_r} \\ &= a_{j_1 m_1} a_{j_2 m_2} \cdots a_{j_r m_r} B_{m_1 m_2 \cdots m_r} \end{aligned}$$

Consequently, $B_{j_1 \cdots j_r}$ does transform as a tensor of rank r as claimed.

Inner product

By the process of forming the outer product of two tensors followed by a contraction, we obtain a new tensor called an *inner product* of the given tensors. If the rank of the two given tensors are respectively r and s then the rank of the new tensor will be $r + s - 2$. Its tensorial nature follows from the fact that an outer product of two tensors is a tensor and the contraction of two indices in a tensor gives a new tensor. In the previous chapters we have reserved the "dot"-symbol for precisely this operation. For instance, a scalar product between two vectors, $\mathbf{a}$ and $\mathbf{b}$ is a inner product of two rank 1 tensors giving a 0 rank tensor (scalar):

$$\begin{aligned} [\mathbf{ab}]_{ij} &= a_i b_j && \text{(Outer product)} \\ \mathbf{a} \cdot \mathbf{b} &= a_i b_i && \text{(Contraction of outer product, setting } j = i) \end{aligned}$$

A rank two tensor, $\mathbf{T}$, operating on a vector $\mathbf{v}$ is a inner product of between a rank two and a rank one tensor giving a rank 1 tensor:

$$[\mathbf{Tv}]_{ijk} = T_{ij}v_k \quad \text{(Outer product)}$$
$$[\mathbf{T}\cdot\mathbf{v}]_i = T_{ij}v_j \quad \text{(Contraction of outer product, setting } k=j\text{)}$$

It is left to an exercise to see that the "dot"-operation defined in chapter 2 for two second rank tensors, $\mathbf{T}\cdot\mathbf{S}$, is an inner product.

In general, any of two indices in the outer product between two tensors can be contracted to define an inner product. For example

$$C_{i_1 i_2 i_4 i_5 j_1 j_3} = A_{i_1 i_2 k i_4 i_5} B_{j_1 k j_3}$$

is a tensor of rank 6 obtained as an inner product between a tensor $A_{i_1 i_2 i_3 i_4 i_5}$ of rank 5 and a tensor $B_{j_1 j_2 j_3}$ of rank 3 by setting $i_3 = j_2$. Tensors of rank higher than 4 are an odditity in the realm of physics, fortunately .

Summary

In summary, all tensorial algebra of any rank can basically be boiled down to these few operations: addition, outer product, permutation, contraction and inner product. Only one more operation is essential to know, namely that of differentiation.

4.3.4 Differentiation

Let us consider a tensor $\mathbf{A}$ of rank r which is a differentiable function of all of the elements of another tensor $\mathbf{B}$ of rank s. The direct notation for this situation would be

$$\mathbf{A} = \mathbf{A}(\mathbf{B}) \qquad (\text{Rank}(\mathbf{A})=\text{r}, \quad \text{Rank}(\mathbf{B})=\text{s})$$

In effect, it implies the existence of 3^r functions, $A_{i_1\cdots i_r}$, each differentiable in each of its 3^s arguments $\underline{B} = (B_{j_1 j_2 \cdots j_s})$,

$$A_{i_1 i_2 \cdots i_r} = A_{i_1 i_2 \cdots i_r}(\ (B_{j_1 j_2 \cdots j_s})\ )$$

Then the partial derivatives

$$C_{i_1 i_2 \cdots i_r, j_1 j_2 \cdots j_s} = \frac{\partial A_{i_1 i_2 \cdots i_r}}{\partial B_{j_1 j_2 \cdots j_s}} \tag{4.22}$$

is itself a tensor of rank $r+s$. As for ordinary derivatives of functions, the components $C_{i_1 i_2 \cdots i_r}$ will in general also be functions of $\underline{B}$,

$$C_{i_1 i_2 \cdots i_r} = C_{i_1 i_2 \cdots i_r}(\ (B_{j_1 j_2 \cdots j_s})\ )$$

The direct notation for taking partial derivatives with respect to a set of tensor components is

$$\mathbf{C}(\mathbf{B}) = \left(\frac{\partial \mathbf{A}}{\partial \mathbf{B}}\right)(\mathbf{B})$$

To demonstrate that $\mathbf{C}$ is a tensor, and so to justify this direct notation in the first place, we will have to look at the transformation properties of its elements. Indeed, we have

$$\begin{aligned} C'_{i_1 i_2 \cdots i_r, j_1 j_2 \cdots j_s}(\underline{B}') &= \left(\frac{\partial A'_{i_1 i_2 \cdots i_r}}{\partial B'_{j_1 j_2 \cdots j_s}}\right)(\underline{B}') \\ &= \frac{\partial B_{l_1 l_2 \cdots l_s}}{\partial B'_{j_1 j_2 \cdots j_s}} \left(\frac{\partial (a_{i_1 k_1} a_{i_2 k_2} \cdots a_{i_r k_r} A_{k_1 k_2 \cdots k_r})}{\partial B_{l_1 l_2 \cdots l_s}}\right)(\underline{B}) \\ &= a_{i_1 k_1} a_{i_2 k_2} \cdots a_{i_r k_r} a_{j_1 l_1} a_{j_2 l_2} \cdots a_{j_s l_s} C_{k_1 k_2 \cdots k_r, l_1 l_2 \cdots l_s}(\underline{B}) \end{aligned} \tag{4.23}$$

In the second step we have used the chain rule for differentiation. The last step follows from

$$B_{l_1 l_2 \cdots l_s} = a_{j_1 l_1} a_{j_2 l_2} \cdots a_{j_s l_s} B'_{j_1 j_2 \cdots j_s}.$$

It is essential that all components of $\mathbf{B}$ are independent and can vary freely without constraints **??**.

From this rule follows the *quotient rule* which states that if the tensor $A_{i_1 i_2 \cdots i_r}$ is a *linear* function of the unconstrained tensor $B_{j_1 j_2 \cdots j_s}$ through the relation

$$A_{i_1 i_2 \cdots i_r} = C_{i_1 i_2 \cdots i_r j_1 j_2 \cdots j_s} B_{j_1 j_2 \cdots j_s} + D_{i_1 i_2 \cdots i_r}$$

then $\mathbf{C}$ is a tensor of rank $r + s$ and consequently $\mathbf{D}$ must also be a tensor of rank r.

4.4 Reflection and pseudotensors

Applying the general tensor transformation formula, Eq. (4.20), to a reflection through the origin of a cartesian coordinate system, $\underline{x}' = -\underline{x}$, we find

$$T'_{i_1 \cdots i_r} = (-1)^r T_{i_1 \cdots i_r} \tag{4.24}$$

There are quantities, notably the Levi-Civita symbol, that do not obey this rule, but acquire an extra minus sign. Such quantities are called *pseudo*-tensors in contradistinction to ordinary or *proper* tensors. Consequently, a set of 3^r quantities, $P_{i_1 i_2 \cdots i_r}$, is called a pseudo-tensor if it transforms according to the rule

$$P'_{i_1 i_2 \cdots i_r} = \det(\underline{A}) a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_r j_r} P_{j_1 j_2 \cdots j_r} \quad \text{Pseudo-tensor,} \tag{4.25}$$

upon the coordinate transformation, Eq. (4.11) For a pure reflection, $a_{ij} = -\delta_{ij}$, so

$$P'_{i_1 i_2 \cdots i_r} = -(-1)^r P_{i 1 i_2 \cdots i_r}.$$

By comparing Eq. (4.25) with Eq. (4.20) one observes that the difference between a tensor and a pseudotensor only appears if the transformation includes a reflection, $\det(\underline{A}) = -1$. A proper tensor can be considered as a real geometrical object, independent of the coordinate system. A proper vector is an "arrow" in space for which the components change sign upon a reflection of the axes of the coordinate system, cf. Eq. (4.24). Direct product of ordinary vectors

are ordinary tensors, changing sign once for each index. An even number of direct products of pseudo-tensors will make an ordinary tensor, whereas an odd number of direct products of pseudo-tensors leads to another pseudo-tensor.

4.4.1 Levi-Civita symbol

If we apply the transformation rule, Eq. (4.20), to the Levi-Civita symbol, ϵ_{ijk} we obtain

$$\epsilon'_{ijk} = a_{il}a_{jm}a_{kn}\epsilon_{lmn}. \tag{4.26}$$

This expression is antisymmetric in all indices ijk [6]. Consequently, it must be proportional to ϵ_{ijk}. To get the constant of proportionality, we should observe the following connection between the determinant of a matrix and the Levi-Civita symbol

$$\det(\underline{A}) = \epsilon_{lmn}a_{1l}a_{2m}a_{3n}, \qquad \underline{A} = (a_{ij}). \tag{4.27}$$

Taking $ijk = 123$ in Eq. (4.26) we get

$$\epsilon'_{123} = \det(\underline{A}) = \det(\underline{A})\epsilon_{123},$$

since $\epsilon_{123} = +1$. Thus the constant of proportionality is $\det(\underline{A})$ and $\epsilon'_{ijk} = \det(\underline{A})\epsilon_{ijk}$. If ϵ'_{ijk} should be identical to ϵ_{ijk} we must multiply the transformation with an extra $\det(\underline{A})$ to account for the possiblity that the transformation involves a reflection, where $\det(\underline{A}) = -1$. The correct transformation law must therefore be

$$\epsilon'_{ijk} = \det(\underline{A})a_{il}a_{jm}a_{kn}\epsilon_{lmn},$$

whence the Levi-Civita symbol is a third rank pseudo-tensor. The transformation rule for pseudo-tensors leaves the Levi-Civita symbol *invariant under all orthogonal transformations*, $\epsilon'_{ijk} = \epsilon_{ijk}$. [7]

The most important application of the Levi-Civita symbol is in the algebraic definition of the cross-product, $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ between two ordinary vectors

$$c_i = [\mathbf{a} \times \mathbf{b}]_i = \epsilon_{ijk}a_j b_k.$$

Since the rhs. is a double inner product between a pseudo-tensor of rank three and two ordinary vectors, the lhs. will be a pseudo-vector. Indeed, the direction of $\mathbf{c}$ depends on the handedness of the coordinate system, as previously mentioned. An equivalent manifestation of its pseudo-vectorial nature is that $\mathbf{c}$ does not change its direction upon an active reflection.

---

[6] For instance,

$\epsilon'_{ikj} = a_{il}a_{km}a_{jn}\epsilon_{lmn} = a_{il}a_{kn}a_{jm}\epsilon_{lnm} = a_{il}a_{jm}a_{kn}\epsilon_{lnm} = -a_{il}a_{jm}a_{kn}\epsilon_{lmn} = -\epsilon'_{ijk}$
The second identity follows from the fact that n and m are bound indices and can therefore be renamed to one and another.

[7] We should not be suprised by the fact that ϵ_{ijk} are unaffected by coordinate transformations, since its definition makes no distinction between the 1,2 and 3 directions.

4.4.2 Manipulation of the Levi-Civita symbol

Two important relations for the Levi-Civita symbol are very useful to derive the myriad of known identities between various vector products. The first is a formula to reduce the product of two Levi-Civita symbols into kronecker deltas

$$\epsilon_{ijk}\epsilon_{lmn} = \det \begin{pmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{pmatrix} \tag{4.28}$$

which after calculating the determinant becomes

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmn} = \quad & \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \\ & -\delta_{in}\delta_{jm}\delta_{kl} - \delta_{jn}\delta_{km}\delta_{il} - \delta_{kn}\delta_{im}\delta_{jl} \end{aligned} \tag{4.29}$$

Due to this relation an arbitrary tensor expression may be reduced to contain zero ϵ-symbols whereas a pseudotensor expression may be reduced to contain only one such symbol. Every time two ϵ's meet in an expression they may be reduced away **??**. This is the secret behind all complicated formules involving vector products. From Eq. (4.29) we may deduce a few more expressions by contraction of indices

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmk} &= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \\ \epsilon_{ijk}\epsilon_{ljk} &= 2\delta_{il} \\ \epsilon_{ijk}\epsilon_{ijk} &= 6 \end{aligned} \tag{4.30}$$

An other important relation for the ϵ-symbol follows from the observation that it is impossible to construct a non-trivial totally antisymmetric symbol with more than three indicies. The reason is that the antisymmetry implies that any component with two equal components must vanish. Hence a non-vanishing component must have all indices different, and since there are only three possible values for an index this is impossible. Consequently, all components of a totally antisymmetric symbol with four indices must vanish. From this follows the rule

$$a_i\epsilon_{jkl} - a_j\epsilon_{ikl} - a_k\epsilon_{jil} - a_l\epsilon_{jki} = 0, \tag{4.31}$$

because the lhs. is antisymmetric in four indices.

Derivation of Eq. (4.28)

To derive Eq. (4.28) one uses two well-known properties of the determinant of a matrix. First, the determinant of a product of matrices equals the product of the determinants of the individual matrices

$$\det(\underline{M} \cdot \underline{N}) = \det(\underline{M}) \det(\underline{N})$$

Secondly, the determinant of the transpose, $\underline{N}^t$, of a matrix, $\underline{N}$, equals the determinant of the matrix

$$\det(\underline{N}^t) = \det(\underline{N})$$

Set

$$\underline{M} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}, \quad \underline{N} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix},$$

then the matrix product $\underline{M} \cdot \underline{N}^t$ becomes

$$\underline{M} \cdot \underline{N}^t = \begin{bmatrix} \mathbf{a}\cdot\mathbf{x} & \mathbf{a}\cdot\mathbf{y} & \mathbf{a}\cdot\mathbf{z} \\ \mathbf{b}\cdot\mathbf{x} & \mathbf{b}\cdot\mathbf{y} & \mathbf{b}\cdot\mathbf{z} \\ \mathbf{c}\cdot\mathbf{x} & \mathbf{c}\cdot\mathbf{y} & \mathbf{c}\cdot\mathbf{z} \end{bmatrix} \tag{4.32}$$

Also, $\det(\underline{M}) = \epsilon_{ijk} a_i b_j c_k$, and $\det(\underline{N}) = \epsilon_{lmn} x_l y_m z_n$ according to Eq. (2.24). Consequently,

$$\det(\underline{M} \cdot \underline{N}^t) = \det(\underline{M}) \det(\underline{N}) = \epsilon_{ijk}\epsilon_{lmn} a_i b_j c_k x_l y_m z_n,$$

which together with Eq. (4.32) gives

$$\epsilon_{ijk}\epsilon_{lmn} a_i b_j c_k x_l y_m z_n = \det \begin{bmatrix} \mathbf{a}\cdot\mathbf{x} & \mathbf{a}\cdot\mathbf{y} & \mathbf{a}\cdot\mathbf{z} \\ \mathbf{b}\cdot\mathbf{x} & \mathbf{b}\cdot\mathbf{y} & \mathbf{b}\cdot\mathbf{z} \\ \mathbf{c}\cdot\mathbf{x} & \mathbf{c}\cdot\mathbf{y} & \mathbf{c}\cdot\mathbf{z} \end{bmatrix}. \tag{4.33}$$

Setting $\mathbf{a} = \mathbf{e}_i$, $\mathbf{b} = \mathbf{e}_j$, $\mathbf{c} = \mathbf{e}_k$, $\mathbf{x} = \mathbf{e}_l$, $\mathbf{y} = \mathbf{e}_m$, $\mathbf{z} = \mathbf{e}_n$ we have for the lhs of Eq. (4.33)

$$\begin{array}{ll} \sum_{i'j'k'l'm'n'} \epsilon_{i'j'k'}\epsilon_{l'm'n'} a_{i'} b_{j'} c_{k'} x_{l'} y_{m'} z_{n'} & = \\ \sum_{i'j'k'l'm'n'} \epsilon_{i'j'k'}\epsilon_{l'm'n'} \delta_{i'i}\delta_{j'j}\delta_{k'k}\delta_{l'l}\delta_{m'm}\delta_{n'n} & = \epsilon_{ijk}\epsilon_{lmn}, \end{array} \tag{4.34}$$

where we for pedagogical reasons have reinserted the involved sums, denoting summation indices with a prime. The above expression equals the lhs of Eq. (4.28). For the rhs of Eq. (4.33) we also recover the rhs. of Eq. (4.28) by noting that each scalar product indeed becomes a kronecker delta, ie. $\mathbf{a}\cdot\mathbf{x} = \mathbf{e}_i \cdot \mathbf{e}_l = \delta_{il}$ etc. This proves Eq. (4.28).

4.5 Tensor fields of any rank

All that has been said about transformation properties of tensors and pseudo-tensors can be converted directly to that of tensor fields and pseudo-tensor fields. One only needs to remember that all the (pseudo) tensor components are now functions of the coordinates. To be specific, upon the cartesian coordinate transformation, Eq. (4.11) the components of a tensor field $\mathbf{T}$ of rank r transforms as

$$T'_{i_1 i_2 \cdots i_r}(\underline{x}') = a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_r j_r} T_{j_1 j_2 \cdots j_r}(\underline{x}) \tag{4.35}$$

and the components of a pseudo tensor field $\mathbf{P}$ of rank r transforms as

$$P'_{i_1 i_2 \cdots i_r}(\underline{x}') = \det(\underline{A}) a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_r j_r} P_{j_1 j_2 \cdots j_r}(\underline{x}) \tag{4.36}$$

The same transformation rules also hold true for a transformation between any two orthonormal bases, for instance from a rectangular to a spherical basis. One must bear in mind, however, than in this case the transformation matrix elements, a_{ij}, will themselves be function of the position, $a_{ij} = a_{ij}(\underline{x})$.

A tensor field is just a particular realization of the more general case considered in section 4.3.4, with a tensor $\mathbf{A}$ being a function of another tensor $\mathbf{B}$, $\mathbf{A} = \mathbf{A}(\mathbf{B})$. Here, $\mathbf{B} = \mathbf{r}$. Therefore, we can apply Eq. (4.22) to demonstrate that derivatives of a tensor field is another tensor field of one higher rank. Although $\mathbf{r}$ is an improper vector, cf. section 2.3, $\mathbf{C} = \frac{\partial \mathbf{T}}{\partial \mathbf{r}}$ will still be a proper tensor. The reason is that in deriving the tensorial nature of $\mathbf{C}$, Eq. (4.23), we have used the tensorial nature of $\mathbf{B}$ only to show that

$$\frac{\partial B_{l_1 l_2 \cdots l_s}}{\partial B'_{j_1 j_2 \cdots j_s}} = a_{j_1 l_1} a_{j_2 l_2} \cdots a_{j_s l_s}$$

However, this also holds true for any improper tensor in cartesian coordinates. Specifically, for the position vector we have $[\mathbf{r}]_{\mathcal{R},l} = x_l$, $[\mathbf{r}]_{\mathcal{R}',j} = x'_j$ and

$$\frac{\partial x_l}{\partial x'_j} = a_{jl}.$$

Consequently, if $\mathbf{T}$ is a tensor field of rank r, $\mathbf{T}(\mathbf{r}) = (T_{i_1 i_2 \cdots i_r}(\mathbf{r}))_{\mathcal{R}}$, then the operation

$$\frac{\partial \mathbf{T}}{\partial \mathbf{r}} = \nabla \mathbf{T}$$

gives a tensor field of rank $r+1$ with the components

$$[\nabla \mathbf{T}]_{i_1 i_2 \cdots i_r, j}(\mathbf{r}) = \big(\partial_j T_{i_1 i_2 \cdots i_r}\big)(\mathbf{r}).$$

Specifically, $\nabla \phi$ is a vector field when $\mathbf{T} = \phi$ is a scalar field, c.f. section 2.6.4, and $\nabla \mathbf{a}$ is a rank two tensor field when $\mathbf{T} = \mathbf{a}$ is a vector, cf. section 3.5. [8]

Divergence

The fact that spatial derivatives of a tensor always leads to a new tensor of one higher rank makes it easy to deduce the type of objects resulting from various vector operations. For instance, if $\mathbf{a}(\underline{x})$ is a vector field then $\nabla \mathbf{a}$ is a tensor field of rank two. By contracting the two indices we therefore obtain a scalar, cf. section 4.3.2

$$[\nabla \mathbf{a}]_{i,i} = \partial_i a_i \qquad \textit{a scalar quantity.}$$

Thus, the *divergence* of the vector field is a scalar.

For a rank two tensor, $\mathbf{T}$, $\nabla \mathbf{T}$ is a rank 3 tensor, and we can perform two different contractions, $a_j = \partial_i T_{ij}$ and $b_i = \partial_j T_{ij}$, yielding the components of two different vectors $\mathbf{a}$ and $\mathbf{b}$. Only if $\mathbf{T}$ is symmetric will $\mathbf{a} = \mathbf{b}$. The direct notation for the two operations are $\nabla \cdot \mathbf{T}$ and $\nabla \cdot \mathbf{T}^t$, respectively.

[8] Note, that the convention in tensor notation is to place the index of the spatial derivative at the end of the tensor components, in conflict with the natural convention arising from the direct notation which for –say a vector $\mathbf{a}$– reads $[\nabla \mathbf{a}]_{ij} = \nabla_i a_j$. However, in tensor notation one always explicitly writes the index to be summed over, so no ambiguities arise in practice.

Curl

The tensor notation for the curl operator $\nabla \times \mathbf{a}$ on a vector field $\mathbf{a}$ involves the Levi-Civita symbol

$$[\nabla \times \mathbf{a}]_i = \epsilon_{ijk} \partial_j a_k.$$

If $\mathbf{a}$ is a polar vector then the rhs will be a double inner product between a pseudo-tensor of rank 3 and two polar vectors yielding an axial (or pseudo) vector field.

Chapter 5

Invariance and symmetries