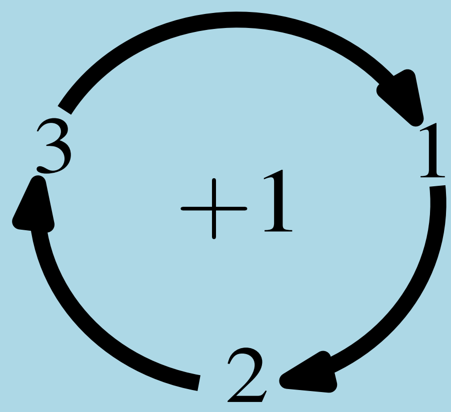
Tensor Calculus

Made Simple

$$\begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}$$



∂_i

$\epsilon_{ijk}\epsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$

$\sum$

∇f

Taha Sochi

Preface

This book is prepared from personal notes and tutorials about tensor calculus at an introductory level. The language and method used in presenting the ideas and techniques of tensor calculus make it very suitable for learning this subject by the beginners who have not been exposed previously to this elegant discipline of mathematics. Yes, some general background in arithmetic, elementary algebra, calculus and linear algebra is needed to understand the book and follow the development of ideas and techniques of tensors. However, we made considerable efforts to reduce this dependency on foreign literature by summarizing the main items needed in this regard to make the book self-contained. The book also contains a number of graphic illustrations to aid the readers and students in their effort to visualize the ideas and understand the abstract concepts. In addition to the graphic illustrations, we used illustrative techniques such as highlighting key terms by boldface fonts.

The book also contains sets of clearly explained exercises which cover most of the materials presented in the book where each set is given in an orderly manner in the end of each chapter. These exercises are designed to provide thorough revisions of the supplied materials and hence they make an essential component of the book and its learning objectives. Therefore, they should not be considered as a decorative accessory to the book. We also populated the text with hyperlinks, for the ebook users, to facilitate referencing and connecting related objects so that the reader can go forth and back with minimum effort and time and without compromising the continuity of reading by interrupting the chain of thoughts.

The book is also furnished with a rather detailed index section which provides access to many keywords and concepts throughout the book. Although this index, like any other index, is not comprehensive, it is supposed to provide access to all the essential terms with particular emphasis on the positions where the enlisted terms have a particular significance in their context such as where they are defined or linked to key ideas. Despite the fact that this index may be of little use to the ebook users who can conduct thorough searches in the book for any term electronically by using the more convenient way of finding words through "Find" function or similar functions in the document viewer, the index is composed with the vision to be like a glossary and hence it serves two purposes. For that reason, the index is used in the exercises as a reference for key terms and concepts of common use in tensor calculus and related keywords from other disciplines of mathematics.

In view of all the above factors, the present text can be used as a textbook or as a reference for an introductory course on tensor algebra and calculus or as a guide for self-studying and learning. I tried to be as clear as possible and to highlight the key issues of the subject at an introductory level in a concise form. I hope I have achieved some success in reaching these objectives for the majority of my target audience.

Finally, I should make a short statement about credits in making this book following the tradition in writing book prefaces. In fact everything in the book is made by the author including all the graphic illustrations, front and back covers, indexing, typesetting, and overall design. However, I should acknowledge the use of the LaTeX typesetting package and the LaTeX based document preparation package LyX for facilitating many things in typesetting and design which cannot be done easily or at all without their versatile and powerful capabilities. I also used the Ipe extensible drawing editor program for making all the graphic illustrations in the book as well as the front and back covers.

Taha Sochi

London, November 2016

Contents

List of Figures

Nomenclature

In the following table, we define some of the common symbols, notations and abbreviations which are used in the book to avoid ambiguity and confusion.

∇	nabla differential operator
∇f	gradient of scalar f
$\nabla \cdot \mathbf{A}$	divergence of vector $\mathbf{A}$
$\nabla \times \mathbf{A}$	curl of vector $\mathbf{A}$
∇^2 or Δ	Laplacian operator
$\perp$	perpendicular to
2D, 3D, nD	two-dimensional, three-dimensional, n-dimensional
det	determinant of matrix
$\mathbf{E}_i$	i^{th} covariant basis vector
$\mathbf{E}^i$	i^{th} contravariant basis vector
Eq./Eqs.	Equation/Equations
h_i	scale factor for i^{th} coordinate in general orthogonal system
iff	if and only if
r, θ, ϕ	coordinates of spherical system in 3D space
tr	trace of matrix
u^1, u^2, u^3	coordinates of general orthogonal system in 3D space
x_1, x_2, x_3	labels of coordinate axes of Cartesian system in 3D space
X_1, X_2, X_3	same as the previous entry
x^1, x^2, x^3	coordinates of general curvilinear system in 3D space
x, y, z	coordinates of points in Cartesian system in 3D space
ρ, ϕ, z	coordinates of cylindrical system in 3D space

Chapter 1

Preliminaries

In this introductory chapter, we provide the reader with a general overview about the historical development of tensor calculus and its role in modern mathematics, science and engineering. We also provide a general set of notes about the notations and conventions which are generally followed in the writing of this book. A general mathematical background about coordinate systems, vector algebra and calculus and matrix algebra is also presented to make the book, to some extent, self-sufficient. Although the general mathematical background section is not comprehensive, it contains essential mathematical terminology and concepts which are needed in the development of the ideas and methods of tensor calculus in the subsequent chapters of the book.

1.1 Historical Overview of Development & Use of Tensor Calculus

First, we should remark that some of the following historical statements represent approximate rather than exact historical facts due to the reality that many of the first hand historical records are missing or not available to the author. Moreover, many ideas, terminology, notation and techniques of tensor calculus, like any other field of knowledge and practice, have been developed gradually over long periods of time although the credit is usually attributed to a few individuals due to their prominence and fame or because they played crucial and distinctive roles in the creation and development of the subject.

It is believed that the word "**tensor**" was coined by **Hamilton** but he used it in a rather different meaning to what is being used for in modern mathematics and science.

The credit for attaching this term to its modern technical meaning, approximately in the late nineteenth century, is usually given to **Voigt**. Apparently, the term "**tensor**" was originally derived from the Latin word "**tensus**" which means tension or stress since one of the first uses of tensors (in whatever meaning) was related to the mathematical description of mechanical stress.

The names "**tensor calculus**" or "**tensor analysis**" have been used to label this subject in its modern form rather recently, probably in the second quarter of the twenties century. The early forms of this subject have been called "**Ricci calculus**" or "**absolute differential calculus**". The latter names may still be found in the modern literature of tensor calculus.

Many mathematicians and scientists have contributed to the development of tensor calculus directly or indirectly. However, numerous components of the modern tensor calculus were not developed as such and for the purpose of the theory of tensors but as parts or byproducts of other disciplines, notably the differential geometry of curves and surfaces. This generally applies prior to the official launch of tensor calculus as an independent branch of mathematics by **Ricci** and **Levi-Civita** who are commonly recognized as the founding fathers of this discipline.

Several concepts and techniques of tensor calculus have been developed by **Gauss** and **Riemann** in the nineteenth century, mainly as part of their efforts to develop and formulate the theory of differential geometry of curves and surfaces. Their contributions are highlighted by the fact that many concepts, methods and equations, which are related directly to tensor calculus or to subjects with a close affinity to tensor calculus, bear their names, e.g. Gaussian coordinates, Gauss curvature tensor, Gauss-Codazzi equations, Riemannian metric tensor, Riemannian manifold, and Riemann-Christoffel curvature tensor.

A major player in the development of tensor calculus is **Christoffel** whose contribution is documented, for instance, in the Christoffel symbols of the first and second kind which infuse throughout the whole subject of tensor calculus and play very essential roles. Also, the above mentioned Riemann-Christoffel curvature tensor is another acknowledgment of his achievements in this respect. **Bianchi** has also contributed a number of important ideas and techniques, such as Bianchi identities, which played an important role in the subsequent development of this subject.

As indicated above, the major role in the creation and development of tensor calculus in its modern style is attributed to **Ricci** and **Levi-Civita** in the end of the nineteenth century and the beginning of the twentieth century where these mathematicians extracted and merged the previous collection of tensor calculus notations and techniques and introduced a number of new ideas, notations and methods and hence they gave the subject its modern face and contemporary style. For this reason, they are usually accredited for the creation of tensor calculus as a standalone mathematical discipline although major components of this subject have been invented by their predecessors as indicated above. The role of **Ricci** and **Levi-Civita** in the creation and renovation of tensor calculus is documented in terms like Ricci calculus, Ricci curvature tensor, Levi-Civita symbol, and Levi-Civita identity.

Many other mathematicians and scientists have made valuable indirect contributions to the subject of tensor calculus by the creation of mathematical ideas which subsequently played important roles in the development of tensor calculus although they are not related directly to tensors. A prominent example of this category is **Kronecker** and his famous Kronecker delta symbol which is embraced in tensor calculus and adapted for extensive use in tensor notation and techniques.

The widespread use of tensor calculus in science has begun with the rise of the **general theory of relativity** which is formulated in the language, methods and techniques of tensor calculus. Tensor calculus has also found essential roles to play in a number of other disciplines with a particular significance in **differential geometry**, **continuum**

mechanics and **fluid dynamics**.

Nowadays, tensor algebra and calculus propagate throughout many branches of mathematics, physical sciences and engineering. The success of tensor calculus is credited to its elegance, power and concision as well as the clarity and eloquence of its notation. Indeed, tensor calculus is not just a collection of mathematical methods and techniques but it is a compact and effective language for communicating ideas and expressing concepts of varying degrees of complexity with rigor and reasonable simplicity.

1.2 General Conventions

In this book, we largely follow certain conventions and general notations; most of which are commonly used in the mathematical literature although they may not be universally approved. In this section, we outline the most important of these conventions and notations. We also give some initial definitions of the most basic terms and concepts in tensor calculus; more thorough technical definitions will follow, if needed, in the forthcoming sections and chapters. It should be remarked that tensor calculus is riddled with conflicting conventions and terminology. In this book we use what we believe to be the most common, clear and useful of all of these.

Scalars are algebraic objects which are uniquely identified by their magnitude (absolute value) and sign ($\pm$), while **vectors** are broadly geometric objects which are uniquely identified by their magnitude (length) and direction in a presumed underlying space. At this early stage in this book, we generically define **tensor** as an organized array of mathematical objects such as numbers or functions.

In generic terms, the **rank** of a tensor signifies the complexity of its structure. Rank-0 tensors are called **scalars** while rank-1 tensors are called **vectors**. Rank-2 tensors may be called **dyads** although dyad, in common use, may be designated to the outer product (see § 3.3) of two vectors and hence it is a special case of rank-2 tensors assuming it meets the requirements of a tensor and hence transforms as a tensor. Like rank-2 tensors, rank-3 tensors may be called **triads**. Similar labels, which are much less common in use, may be attached to higher rank tensors; however, none of these will be used in the present book. More generic names for higher rank tensors, such as **polyad**, are also in use.

We may use the term "**tensor**" to mean tensors of all ranks including scalars (rank-0) and vectors (rank-1). We may also use this term as opposite to scalar and vector, i.e. tensor of rank-n where $n > 1$. In almost all cases, the meaning should be obvious from the context. It should be remarked that in the present book all tensors of all ranks and types are assumed to be **real** quantities, i.e. they have real rather than complex components.

Due to its introductory nature, the present book is largely based on assuming an underlying **orthonormal Cartesian** coordinate system. However, certain parts of the book are based on other types of coordinate system; in these cases this is stated explicitly or made clear by the notation and context. As will be outlined later (see § Orthonormal Cartesian Coordinate System), we mean by "**orthonormal**" a system with mutually perpendicular and uniformly scaled axes with a unit basis vector set. We may also use "**rectangular**" for a similar meaning or to exclude **oblique** coordinate systems[1] which may also be labeled by some as Cartesian. Some of the statements in this book in which these terms are used may not strictly require these conditions but we add these terms to focus the attention on the type of the coordinate system which these statements are made about.

To label **scalars**, we use non-indexed lower case light face italic Latin letters (e.g. f and h), while for **vectors**, we use non-indexed lower or upper case bold face non-italic Latin letters (e.g. $\mathbf{a}$ and $\mathbf{A}$). The exception to this is the basis vectors where indexed bold face lower or upper case non-italic symbols (e.g. $\mathbf{e}_1$ and $\mathbf{E}^i$) are used. However, there should be no confusion or ambiguity about the meaning of any one of these symbols. As for **tensors** of rank > 1, non-indexed upper case bold face non-italic Latin letters (e.g. $\mathbf{A}$ and $\mathbf{B}$) are

[1] Oblique coordinate systems are usually characterized by similar features as orthonormal Cartesian systems but with non-perpendicular axes.

used.[2]

Indexed light face italic Latin symbols (e.g. a_i and B_i^{jk}) are used in this book to denote **tensors of rank** > 0 in their explicit tensor form, i.e. **index notation**. Such symbols may also be used to denote the components of these tensors. The meaning is usually transparent and can be identified from the context if it is not declared explicitly. Tensor indices in this book are lower case Latin letters, usually taken from the middle of the Latin alphabet like (i, j, k). We also use numbered indices like $(i_1, i_2, \ldots, i_k)$ when the number of tensor indices is variable. Numbers are also used as indices (e.g. ϵ_{12}) in some occasions for obvious purposes such as making statements about particular components.

Mathematical **identities** and **definitions** are generally denoted in the mathematical literature by using the symbol "$\equiv$". However, for simplicity we use the equality sign "$=$" to mark identities and mathematical definitions as well as normal equalities.

We use **vertical bars** (i.e. $|\cdot|$) to symbolize determinants and **square brackets** (i.e. $[\cdot]$) to symbolize matrices. This applies where these symbols contain arrays of objects; otherwise they have their normal meaning according to the context, e.g. bars embracing vectors mean modulus of the vector. We normally use **indexed square brackets** (e.g. $[\mathbf{A}]_i$ and $[\nabla f]_i$) to denote the i^{th} component of vectors in their symbolic or vector notation. For tensors of higher rank, more than one index is needed to denote their components, e.g. $[\mathbf{A}]_{ij}$ for the ij^{th} component of a rank-2 tensor.

Partial derivative symbol with a subscript index (e.g. ∂_i) is frequently used to denote the i^{th} component of the **gradient operator nabla** ∇:[3]

$$\partial_i = \nabla_i = \frac{\partial}{\partial x_i} \tag{1}$$

A **comma** preceding a subscript index (e.g. $_{,i}$) is also used to denote **partial differen-**

[2] Since matrices in this book are supposed to represent rank-2 tensors, they also follow the rules of labeling tensors symbolically by using non-indexed upper case bold face non-italic Latin letters.

[3] This mainly applies in this book to Cartesian coordinate systems.

tiation with respect to the i^{th} spatial coordinate, mainly in Cartesian systems, e.g.

$$A_{,i} = \frac{\partial A}{\partial x_i} \tag{2}$$

Partial derivative symbol with a spatial subscript, rather than an index, is used to denote partial differentiation with respect to that spatial variable. For instance:

$$\partial_r = \nabla_r = \frac{\partial}{\partial r} \tag{3}$$

is used for the partial derivative with respect to the radial coordinate r in spherical coordinate systems identified by the spatial variables (r, θ, ϕ).[4]

Partial derivative symbol with repeated double index is used to denote the **Laplacian operator**:

$$\partial_{ii} = \partial_i \partial_i = \nabla^2 = \Delta \tag{4}$$

The notation is not affected by using repeated double index other than i (e.g. ∂_{jj} or ∂_{kk}). The following notations:

$$\partial_{ii}^2 \qquad \partial^2 \qquad \partial_i \partial^i \tag{5}$$

are also used in the literature of tensor calculus to symbolize the Laplacian operator. However, these notations will not be used in the present book.

We follow the common convention of using a **subscript semicolon** preceding a subscript index (e.g. $A_{kl;i}$) to symbolize the operation of **covariant differentiation** with respect to the i^{th} coordinate (see § 7). The semicolon notation may also be attached to the normal differential operators for the same purpose, e.g. $\nabla_{;i}$ or $\partial_{;i}$ to indicate covariant differentiation with respect to the variable indexed by i.

[4] As indicated, in notations like ∂_r the subscript is used as a label rather than an index.

Finally, all transformation equations in the present book are assumed to be **continuous** and **real**, and all derivatives are **continuous** in their domain of variables. Based on the continuity condition of the differentiable quantities, the individual differential operators in the second (and higher) order partial derivatives with respect to different indices are **commutative**, that is:

$$\partial_i \partial_j = \partial_j \partial_i \tag{6}$$

1.3 General Mathematical Background

In this section, we provide a general mathematical background which is largely required to understand the forthcoming materials about tensors and their algebra and calculus. This mathematical background is intended to provide the essential amount needed for the book to be self-contained, to some degree, but it is not comprehensive. Also some concepts and techniques discussed in this section require more elementary materials which should be obtained from textbooks at more basic levels.

1.3.1 Coordinate Systems

In generic terms, a coordinate system is a mathematical device used to identify the location of points in a given space. In tensor calculus, a coordinate system is needed to define non-scalar tensors in a specific form and identify their components in reference to the basis set of the system. Hence, non-scalar tensors require a predefined coordinate system to be fully identified.[5]

There are many types of coordinate system, the most common ones are: the orthonormal Cartesian, the cylindrical and the spherical. A 2D version of the cylindrical system is the

[5] There are special tensors of numerical nature, such as the Kronecker and permutation tensors, which do not require a particular coordinate system for their full and unambiguous identification since their components are invariant under coordinate transformations (the reader is referred to Chapter 4 for details). However, this issue may be debated.

plane polar system. The most general type of coordinate system is the general curvilinear. A subset of the latter is the orthogonal curvilinear. These types of coordinate system are briefly investigated in the following subsections.

A. Orthonormal Cartesian Coordinate System

This is the simplest and the most commonly used coordinate system. It consists, in its simplest form, of three mutually orthogonal straight **axes** that meet at a common point called the **origin of coordinates** O. The three axes, assuming a 3D space, are scaled uniformly and hence they all have the same unit length. Each axis has a unit vector oriented along the positive direction of that axis.[6] These three unit vectors are called the **basis vectors** or the bases of the system. These basis vectors are constant in magnitude and direction throughout the system.[7] This system with its basis vectors ($\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$) is depicted in Figure 1.

The three axes, as well as the basis vectors, are usually labeled according to the **right hand rule**, that is if the index finger of the right hand is pointing in the positive direction of the first axis and its middle finger is pointing in the positive direction of the second axis then the thumb will be pointing in the positive direction of the third axis.

B. Cylindrical Coordinate System

The cylindrical coordinate system is defined by three parameters: ρ, ϕ and z which range over: $0 \le \rho < \infty$, $0 \le \phi < 2\pi$ and $-\infty < z < \infty$. These parameters identify the coordinates of a point P in a 3D space where ρ represents the **perpendicular distance** from the point to the x_3-axis of a corresponding rectangular Cartesian system, ϕ represents the **angle** between the x_1-axis and the line connecting the origin of coordinates O to the

[6] As indicated before, these features are what qualify a Cartesian system to be described as orthonormal.

[7] In fact, the basis vectors are constant only in rectangular Cartesian and oblique Cartesian coordinate systems. As indicated before, the oblique Cartesian systems are the same as the rectangular Cartesian but with the exception that their axes are not mutually orthogonal. Also, labeling the oblique as Cartesian may be controversial.

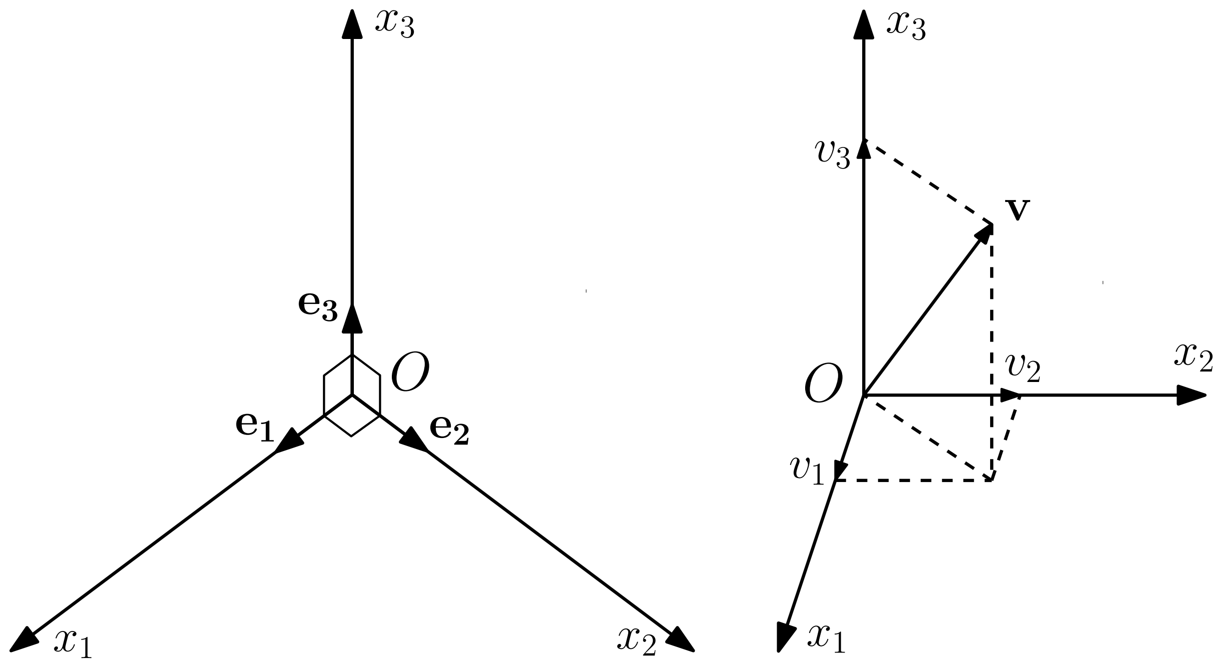


Figure 1: Orthonormal right-handed Cartesian coordinate system and its basis vectors $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$ in a 3D space (left frame) with the components of a vector $\mathbf{v}$ in this system (right frame).

perpendicular projection of the point on the x_1-x_2 plane of the corresponding Cartesian system, and z is the same as the third coordinate of the point in the reference Cartesian system. The sense of the angle ϕ is given by the right hand twist rule, that is if the fingers of the right hand curl in the sense of rotation from the x_1-axis towards the line of projection, then the thumb will be pointing in the positive direction of the x_3-axis.

The **basis vectors** of the cylindrical system ($\mathbf{e}_\rho, \mathbf{e}_\phi$ and $\mathbf{e}_z$) are pointing in the direction of increasing ρ, ϕ and z respectively. Hence, while $\mathbf{e}_z$ is constant in magnitude and direction throughout the system, $\mathbf{e}_\rho$ and $\mathbf{e}_\phi$ are **coordinate-dependent** as they vary in direction from point to point. All these basis vectors are **mutually perpendicular** and they are defined to be of **unit length**. Figure 2 is a graphical illustration of the cylindrical coordinate system and its basis vectors with a corresponding reference Cartesian system in a standard position.

The **transformation** from the Cartesian coordinates (x, y, z) of a particular point in

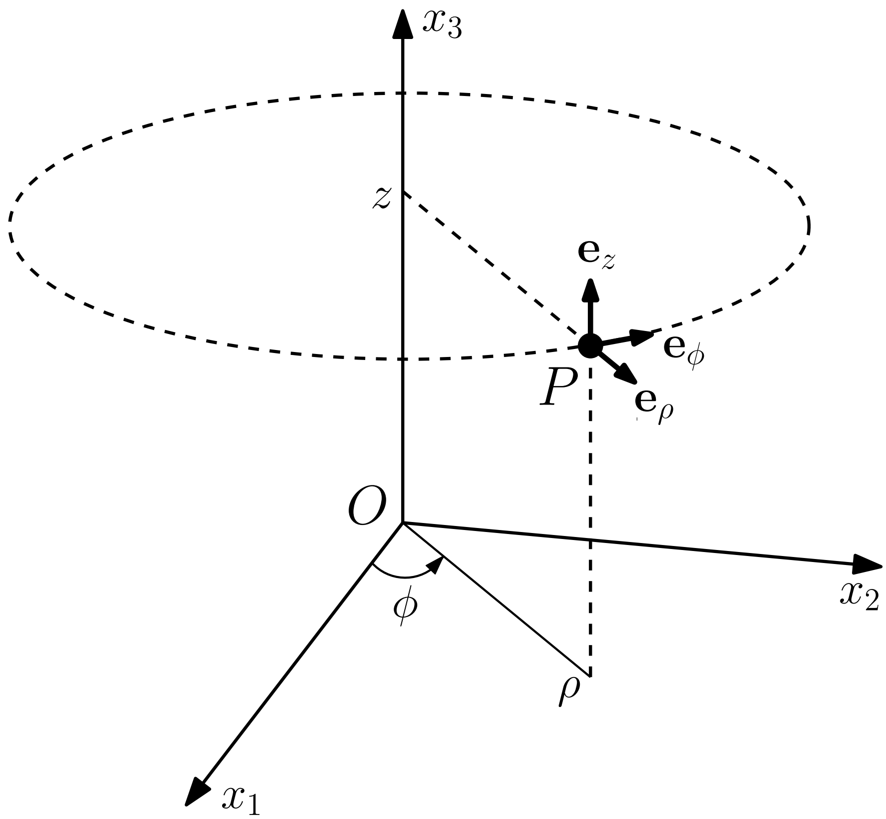


Figure 2: Cylindrical coordinate system, superimposed on a rectangular Cartesian system in a standard position, and its basis vectors $\mathbf{e}_\rho, \mathbf{e}_\phi$ and $\mathbf{e}_z$ in a 3D space. The point P in the figure is identified simultaneously by (x, y, z) coordinates in the Cartesian system and by (ρ, ϕ, z) coordinates in the cylindrical system where these coordinates are related through the two sets of Eqs. 7 and 8.

the space to the cylindrical coordinates (ρ, ϕ, z) of that point, where the two systems are in a standard position, is performed through the following equations:[8]

$$\rho = \sqrt{x^2 + y^2} \qquad \phi = \arctan\left(\frac{y}{x}\right) \qquad z = z \tag{7}$$

while the **opposite transformation** from the cylindrical to the Cartesian coordinates is

[8] In the second equation, $\arctan\left(\frac{y}{x}\right)$ should be selected consistent with the signs of x and y to be in the right quadrant.

performed by the following equations:

$$x = \rho \cos \phi \qquad y = \rho \sin \phi \qquad z = z \tag{8}$$

C. Plane Polar Coordinate System

The plane polar coordinate system is the same as the cylindrical coordinate system with the absence of the z coordinate, and hence all the equations of the polar system are obtained by setting $z = 0$ in the equations of the cylindrical coordinate system, that is:

$$\rho = \sqrt{x^2 + y^2} \qquad \phi = \arctan\left(\frac{y}{x}\right) \tag{9}$$

and

$$x = \rho \cos \phi \qquad y = \rho \sin \phi \tag{10}$$

This system is illustrated graphically in Figure 3 where the point P is located by (ρ, ϕ) in the polar system and by (x, y) in the corresponding 2D Cartesian system.

D. Spherical Coordinate System

The spherical coordinate system is defined by three parameters: r, θ and ϕ which range over: $0 \leq r < \infty$, $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$. These parameters identify the coordinates of a point P in a 3D space where r represents the **distance** from the origin of coordinates O to P, θ is the **angle** from the positive x_3-axis of the corresponding Cartesian system to the line connecting the origin of coordinates O to P, and ϕ is the same as in the cylindrical coordinate system.

The **basis vectors** of the spherical system ($\mathbf{e}_r, \mathbf{e}_\theta$ and $\mathbf{e}_\phi$) are pointing in the direction of increasing r, θ and ϕ respectively. Hence, all these basis vectors are **coordinate-dependent** as they vary in direction from point to point. All these basis vectors are

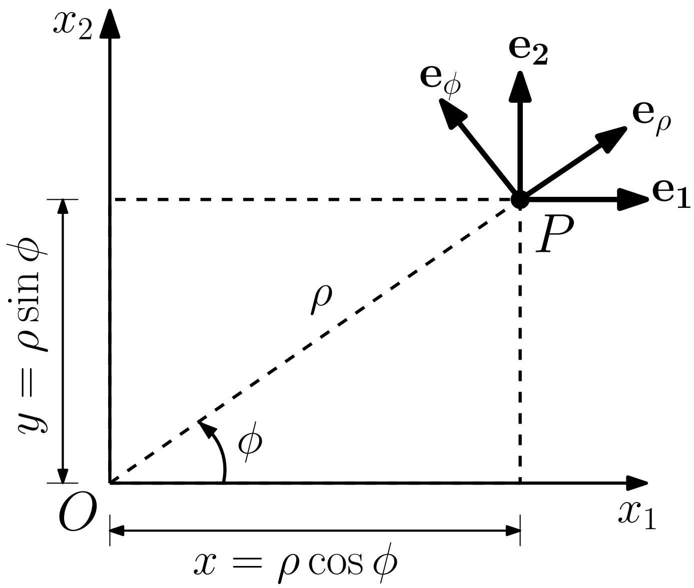


Figure 3: Plane polar coordinate system, superimposed on a 2D rectangular Cartesian system in a standard position, and their basis vectors $\mathbf{e}_\rho$ and $\mathbf{e}_\phi$ for the polar system and $\mathbf{e}_1$ and $\mathbf{e}_2$ for the Cartesian system. The point P in the figure is identified simultaneously by (x, y) coordinates in the Cartesian system and by (ρ, ϕ) coordinates in the polar system where these coordinates are related through the two sets of Eqs. 9 and 10.

mutually perpendicular and they are defined to be of **unit length**. Figure 4 is a graphical illustration of the spherical coordinate system and its basis vectors with a corresponding reference rectangular Cartesian system in a standard position.

The **transformation** from the Cartesian coordinates (x, y, z) of a particular point in the space to the spherical coordinates (r, θ, ϕ) of that point, where the two systems are in a standard position, is performed by the following equations:[9]

$$r = \sqrt{x^2 + y^2 + z^2} \qquad \theta = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \qquad \phi = \arctan\left(\frac{y}{x}\right) \quad (11)$$

[9] Again, $\arctan\left(\frac{y}{x}\right)$ in the third equation should be selected consistent with the signs of x and y.

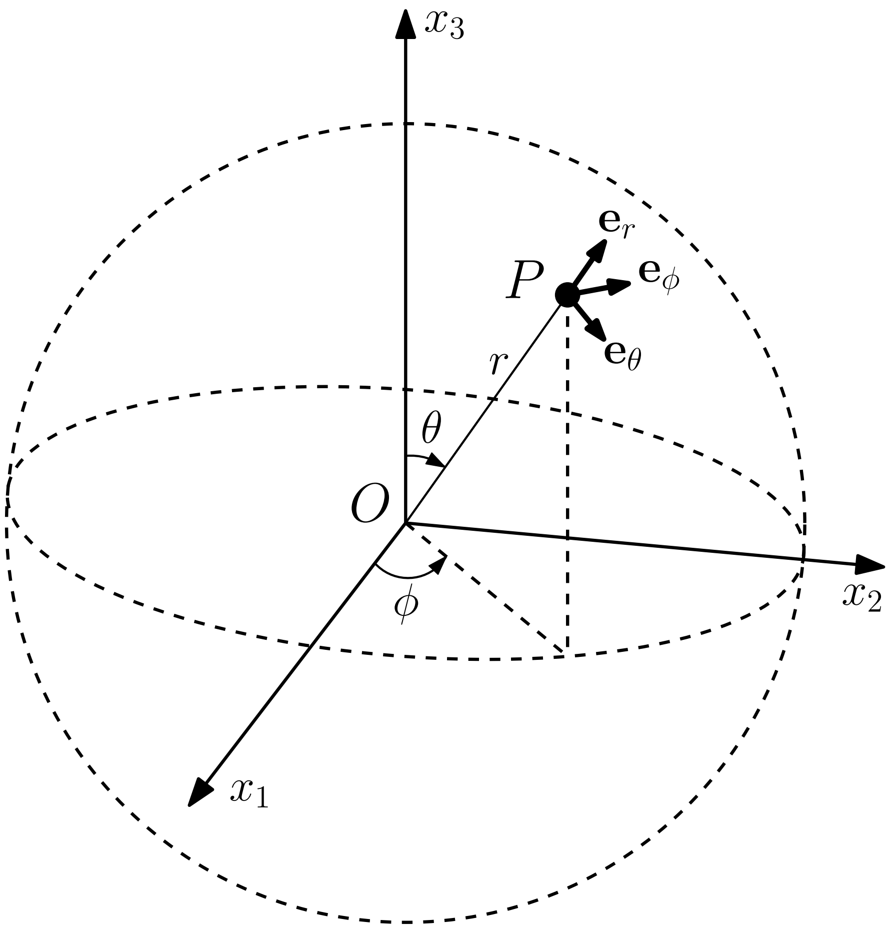


Figure 4: Spherical coordinate system, superimposed on a rectangular Cartesian system in a standard position, and its basis vectors $\mathbf{e}_r, \mathbf{e}_\theta$ and $\mathbf{e}_\phi$ in a 3D space. The point P in the figure is identified simultaneously by (x, y, z) coordinates in the Cartesian system and by (r, θ, ϕ) coordinates in the spherical system where these coordinates are related through the two sets of Eqs. 11 and 12.

while the **opposite transformation** from the spherical to the Cartesian coordinates is performed by the following equations:

$$x = r \sin\theta \cos\phi \qquad y = r \sin\theta \sin\phi \qquad z = r \cos\theta \tag{12}$$

E. General Curvilinear Coordinate System

The general curvilinear coordinate system is characterized by having coordinate axes which are curved in general. Also, its **basis vectors** are generally **position-dependent** and hence they are variable in magnitude and direction throughout the system. Consequently, the basis vectors are **not** necessarily of **unit length** or **mutually orthogonal**. A graphic demonstration of the general curvilinear coordinate system at a particular point in the space with its covariant basis vectors (see § 2.6.1) is shown in Figure 5.

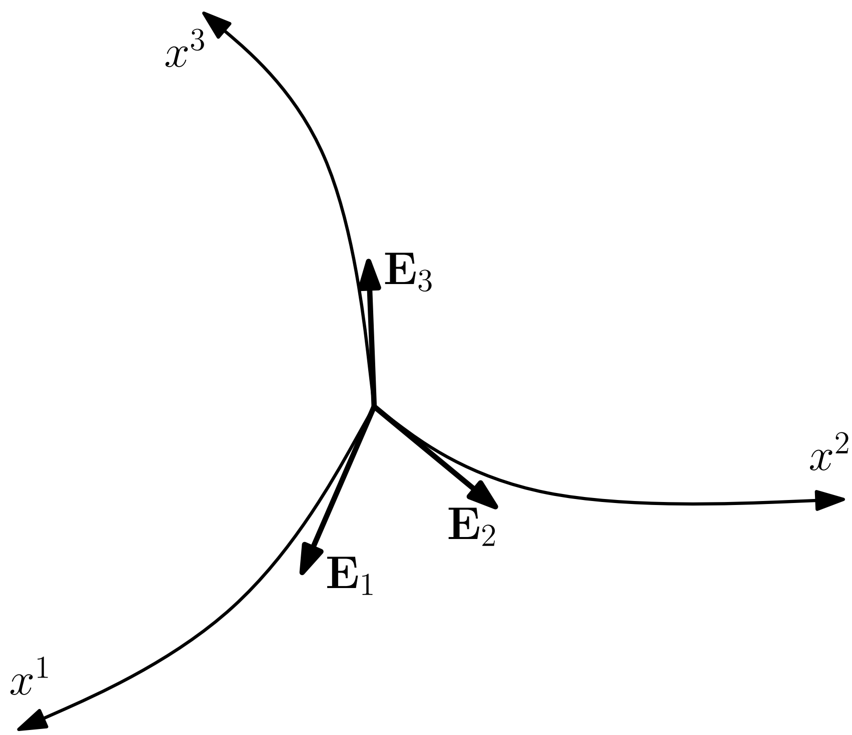


Figure 5: General curvilinear coordinate system and its covariant basis vectors $\mathbf{E}_1, \mathbf{E}_2$ and $\mathbf{E}_3$ (see § 2.6.1) in a 3D space, where x^1, x^2 and x^3 are the labels of the coordinates.

F. General Orthogonal Curvilinear Coordinate System

The general orthogonal curvilinear coordinate system is a **subset** of the general curvilinear coordinate system as described above. It is distinguished from the other subsets of the general curvilinear system by having **coordinate axes** and **basis vectors** which are **mutually orthogonal** throughout the system. The cylindrical and spherical systems are examples of orthogonal curvilinear coordinate systems.

1.3.2 Vector Algebra and Calculus

This subsection provides a short introduction to vector algebra and calculus, a subject that is closely related to tensor calculus. In fact many ideas and methods of tensor calculus find their precursors and roots in vector algebra and calculus and hence many concepts, techniques and notations of the former can be viewed as extensions and generalizations of their counterparts in the latter.

A. Dot Product of Vectors

The dot product, or **scalar product**, of two vectors is a **scalar** quantity which has two interpretations: geometric and algebraic. **Geometrically**, the dot product of two vectors $\mathbf{a}$ and $\mathbf{b}$ can be interpreted as the **projection** of $\mathbf{a}$ onto $\mathbf{b}$ times the **length** of $\mathbf{b}$, or as the projection of $\mathbf{b}$ onto $\mathbf{a}$ times the length of $\mathbf{a}$, as demonstrated in Figure 6. In both cases, the dot product is obtained by taking the product of the **length** of the two vectors involved times the **cosine** of the angle between them when their tails are made to coincide, that is:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos\theta \tag{13}$$

where the **dot** between $\mathbf{a}$ and $\mathbf{b}$ on the left hand side of the equation stands for the dot product operation, $0 \le \theta \le \pi$ is the angle between the two vectors and the bars notation means the **modulus** or the length of the vector.

Algebraically, the dot product is the sum of the products of the **corresponding components** of the two vectors, that is:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \tag{14}$$

where a_i and b_j $(i, j = 1, 2, 3)$ are the components of $\mathbf{a}$ and $\mathbf{b}$ respectively. Here, we are assuming an orthonormal Cartesian system in a 3D space; the formula can be easily

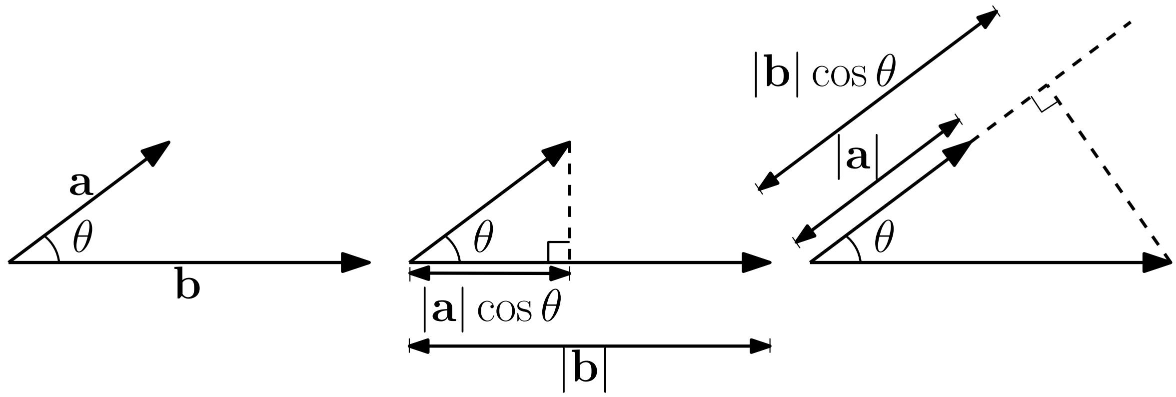


Figure 6: Demonstration of the geometric interpretation of the dot product of two vectors **a** and **b** (left frame) as the projection of **a** onto **b** times the length of **b** (middle frame) or as the projection of **b** onto **a** times the length of **a** (right frame).

extended to an nD space, that is:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i \tag{15}$$

From Eq. 13, it is obvious that the dot product is **positive** when $0 \leq \theta < \frac{\pi}{2}$, **zero** when $\theta = \frac{\pi}{2}$ (i.e. the two vectors are orthogonal), and **negative** when $\frac{\pi}{2} < \theta \leq \pi$. The **magnitude** of the dot product is **equal** to the product of the lengths of the two vectors when they have the same orientation (i.e. parallel or anti-parallel). Based on the above given facts, the dot product is **commutative**, that is:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} \tag{16}$$

B. Cross Product of Vectors

Geometrically, the cross product, or **vector product**, of two vectors, **a** and **b**, is a **vector** whose **length** is equal to the **area** of the parallelogram defined by the two vectors as its two main sides when their tails coincide and whose orientation is **perpendicular** to the plane of the parallelogram with a direction defined by the **right hand rule** as demonstrated in Figure 7. Hence the cross product of two vectors **a** and **b** is given by:

$$\mathbf{a} \times \mathbf{b} = (|\mathbf{a}|\,|\mathbf{b}|\sin\theta)\,\mathbf{n} \tag{17}$$

where $0 \le \theta \le \pi$ is the angle between the two vectors when their tails coincide and **n** is a unit vector perpendicular to the plane containing **a** and **b** and is directed according to the right hand rule.

From Eq. 17, it can be seen that the cross product vector is **zero** when $\theta = 0$ or $\theta = \pi$, i.e. when the two vectors are parallel or anti-parallel respectively. Also, the length of the cross product vector is **equal** to the product of the lengths of the two vectors when the two vectors are orthogonal (i.e. $\theta = \frac{\pi}{2}$).

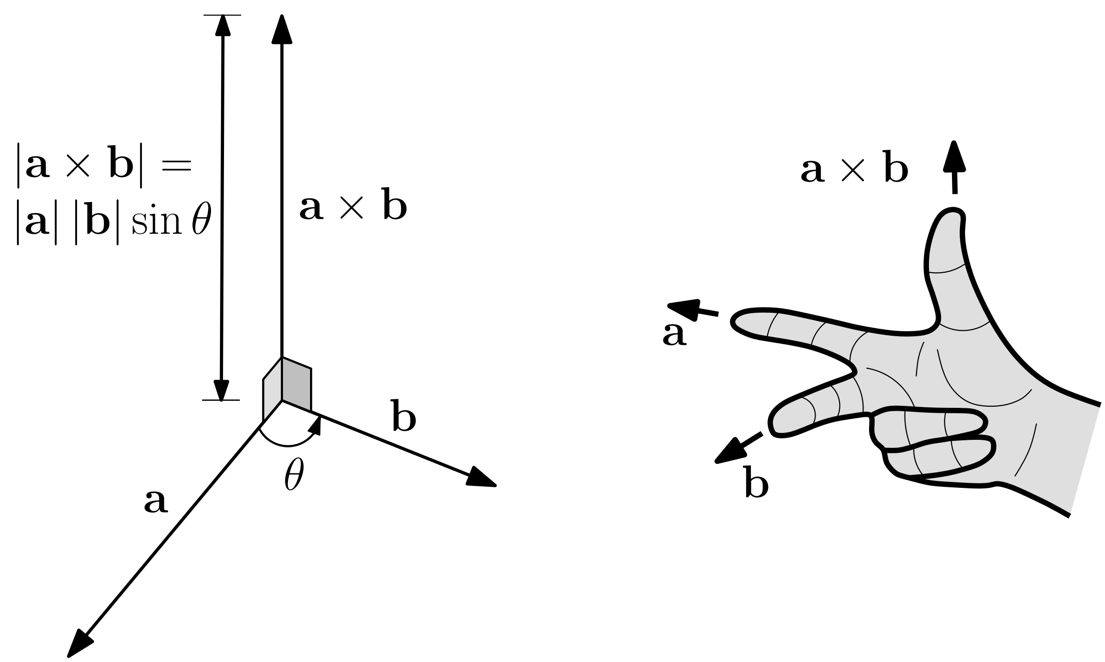


Figure 7: Graphical demonstration of the cross product of two vectors **a** and **b** (left frame) with the right hand rule (right frame).

Algebraically, the cross product of two vectors **a** and **b** is expressed by the following **determinant** (see § Determinant of Matrix) where the determinant is expanded along its first row, that is:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2)\,\mathbf{e}_1 + (a_3 b_1 - a_1 b_3)\,\mathbf{e}_2 + (a_1 b_2 - a_2 b_1)\,\mathbf{e}_3 \qquad (18)$$

Here, we are assuming an orthonormal Cartesian coordinate system in a 3D space with a basis vector set $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$.

Based on the above given facts, since the direction is determined by the right hand rule the cross product is **anti-commutative**, that is:

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a} \qquad (19)$$

C. Scalar Triple Product of Vectors

The scalar triple product of three vectors ($\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$) is a **scalar** quantity defined by the expression:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \qquad (20)$$

where the dot and multiplication symbols stand respectively for the dot and cross product operations of two vectors as defined above. Hence, the scalar triple product is defined **geometrically** by:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = |\mathbf{a}|\,|\mathbf{b}|\,|\mathbf{c}| \sin\phi \cos\theta \qquad (21)$$

where ϕ is the angle between $\mathbf{b}$ and $\mathbf{c}$ while θ is the angle between $\mathbf{a}$ and $\mathbf{b} \times \mathbf{c}$. The scalar triple product is illustrated graphically in Figure 8. As there is no meaning of a cross product operation between a scalar and a vector, the **parentheses** in the above equation are redundant although they provide a clearer and more clean notation.

Now, since $|\mathbf{b} \times \mathbf{c}|$ $(= |\mathbf{b}|\,|\mathbf{c}| \sin\phi)$ is equal to the **area** of the parallelogram whose two

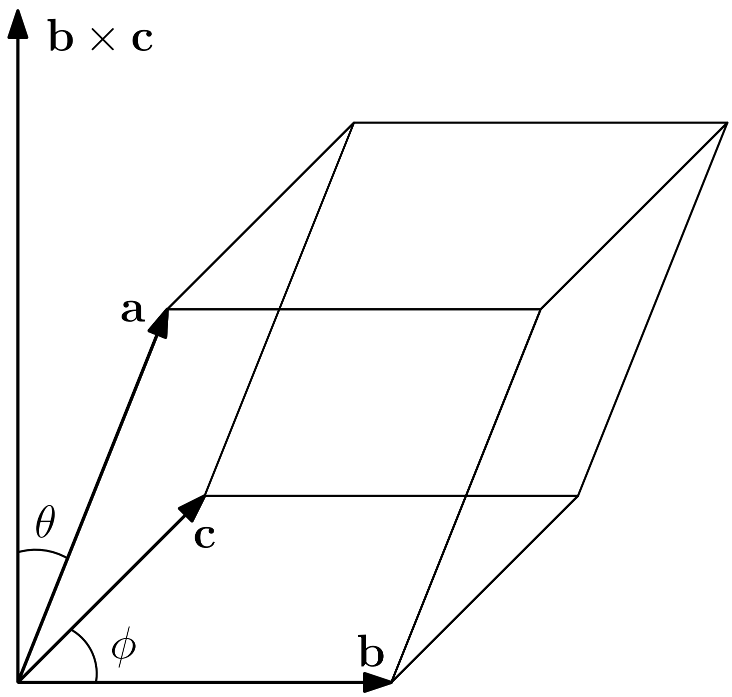


Figure 8: Graphic illustration of the scalar triple product of three vectors **a**, **b** and **c**. The magnitude of this product is equal to the volume of the seen parallelepiped.

main sides are **b** and **c**, while $|\mathbf{a}|\cos\theta$ represents the projection of **a** onto the orientation of $\mathbf{b} \times \mathbf{c}$ and hence it is equal to the **height** of the parallelepiped (refer to Figure 8), the **magnitude** of the scalar triple product is equal to the **volume** of the parallelepiped whose three main sides are $\mathbf{a}, \mathbf{b}$ and **c** while its **sign** is positive or negative depending, respectively, on whether the vectors $\mathbf{a}, \mathbf{b}$ and **c** form a right-handed or left-handed system.

The scalar triple product is **invariant** to a **cyclic permutation** of the symbols of the three vectors involved, that is:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) \tag{22}$$

It is also **invariant** to an **exchange** of the dot and cross product symbols, that is:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \tag{23}$$

Hence, from the three possibilities of the first invariance with the two possibilities of the second invariance, we have **six** equal expressions for the scalar triple product of three vectors.[10] The **other six** possibilities of the scalar triple product of three vectors, which are obtained from the first six possibilities with the opposite cyclic permutations, are also equal to each other for the same reason. However, they are equal in magnitude to the first six possibilities but are different in sign.[11]

From the above interpretation of the scalar triple product as the **signed volume** of the parallelepiped formed by the three vectors, it is obvious that this product is **zero** when the three vectors are **coplanar**. This, of course, includes the possibility of being **collinear**.

The scalar triple product of three vectors is also defined **algebraically** as the **determinant** (refer to § Determinant of Matrix) of the matrix formed by the components of the three vectors as its rows or columns in the given order, that is:

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \det(\mathbf{a},\mathbf{b},\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \tag{24}$$

D. Vector Triple Product of Vectors

The vector triple product of three vectors ($\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$) is a **vector** quantity defined by the following expressions:

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) \qquad \text{or} \qquad (\mathbf{a}\times\mathbf{b})\times\mathbf{c} \tag{25}$$

where the multiplication symbols stand for the cross product operation of two vectors as

[10] This equality can be explained by the fact that the magnitude of the scalar triple product is equal to the volume of the parallelepiped whereas the cyclic permutation preserves the sign of these six possibilities.

[11] The stated facts about the other six possibilities can be explained by the two invariances (as explained in the previous footnote) plus the fact that the cross product operation is anti-commutative.

defined previously. Vector triple product is **not associative**, so in general we have:[12]

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \tag{26}$$

E. Differential Operations of nabla Operator[13]

There are many differential operations that can be performed on scalar and vector fields.[14] Here, we define the most important and widely used of these operations that involve the **nabla** ∇ differential operator, that is the **gradient**, **divergence** and **curl**. We also define the **Laplacian** differential operator which is based on a combination of the gradient and divergence operations. The definitions here are given in Cartesian coordinates only. More definitions of these operations and operators will be given in § 5.3 in terms of tensor notation for the Cartesian, as well as the equivalent definitions for some non-Cartesian coordinate systems.

The differential vector operator **nabla** ∇ is defined in rectangular Cartesian coordinate systems by the following expression:

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k} \tag{27}$$

where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors in the x, y and z directions respectively.

The **gradient of a scalar field** $f(x, y, z)$ is a **vector** defined, in Cartesian coordinate

[12] Associativity here should be understood in its context as stated by the inequality.

[13] Although we generally use (x, y, z) for the Cartesian coordinates of a particular point in the space while we use (x_1, x_2, x_3) to label the axes and the coordinates of the Cartesian system in general, in this subsection we use (x, y, z), instead of (x_1, x_2, x_3), because it is more commonly used in vector algebra and calculus and is notationally clearer, especially at this level and at this stage in the book. We also label the basis vectors of the Cartesian system with $(\mathbf{i}, \mathbf{j}, \mathbf{k})$, instead of $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, for similar reasons.

[14] As we will see later in the book, similar differential operations can also be defined and performed on higher rank tensor fields.

systems, by:

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} \tag{28}$$

Geometrically, the gradient of a scalar field $f(x, y, z)$, at any point in the space where the field is defined, is a **vector normal** to the surface $f(x, y, z) = \text{constant}$ (refer to Figure 9) pointing in the direction of the fastest increase in the field at that point.

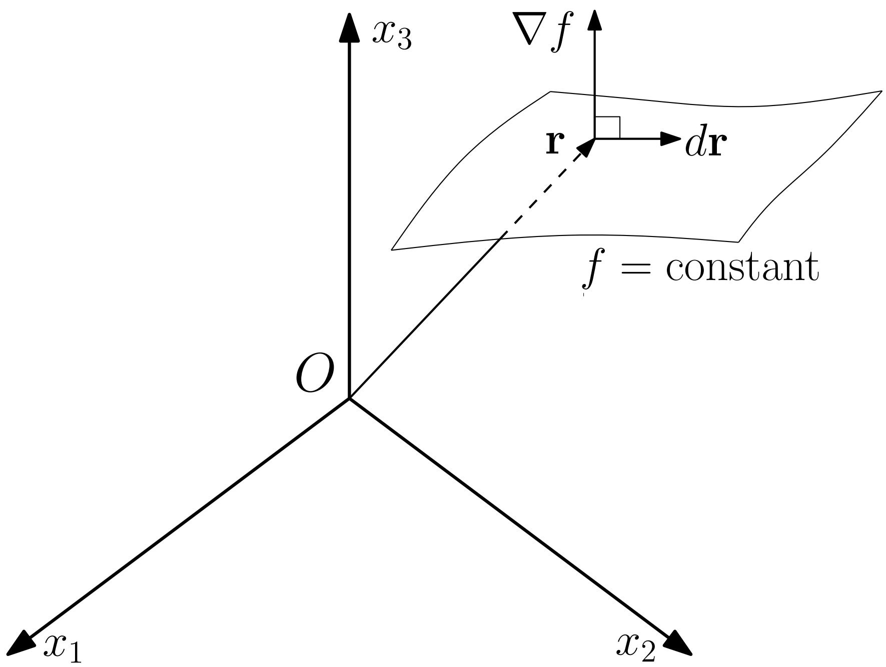


Figure 9: The gradient of a scalar field $f(x, y, z)$ as a vector normal to the surface $f(x, y, z) = \text{constant}$ pointing in the direction of the fastest increase in the field at that point.

The gradient operation is **distributive** but **not commutative** or **associative**, that is:

$$\nabla (f + h) = \nabla f + \nabla h \tag{29}$$

$$\nabla f \neq f \nabla \tag{30}$$

$$(\nabla f)\, h \neq \nabla (f h) \tag{31}$$

where f and h are differentiable scalar functions of position.

The **divergence of a vector field** $\mathbf{v}(x, y, z)$ is a **scalar** quantity defined as the dot product of the nabla operator with the vector. Hence, in Cartesian coordinate systems it is given by:

$$\nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \tag{32}$$

where v_x, v_y and v_z are the components of $\mathbf{v}$ in the x, y and z directions respectively. In broad terms, the physical significance of the divergence of a vector field is that it is a measure of how much the field **diverges** or **converges** at a particular point in the space where the field is defined. When the divergence of a vector field is identically zero, the field is called **solenoidal**.

The divergence operation is **distributive** but **not commutative** or **associative**, that is:

$$\nabla \cdot (\mathbf{A} + \mathbf{B}) = \nabla \cdot \mathbf{A} + \nabla \cdot \mathbf{B} \tag{33}$$

$$\nabla \cdot \mathbf{A} \neq \mathbf{A} \cdot \nabla \tag{34}$$

$$\nabla \cdot (f\mathbf{A}) \neq \nabla f \cdot \mathbf{A} \tag{35}$$

where $\mathbf{A}$ and $\mathbf{B}$ are differentiable vector functions of position.

The **curl of a vector field** $\mathbf{v}(x, y, z)$ is a **vector** defined as the cross product of the nabla operator with the vector. Hence, in Cartesian coordinate systems it is given by (refer to § Cross Product of Vectors):

$$\nabla \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_x & v_y & v_z \end{vmatrix} = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \mathbf{k} \tag{36}$$

Broadly speaking, the curl of a vector field is a quantitative measure of the **circulation** or **rotation** of the field at a given point in the space where the field is defined. When the curl of a vector field vanishes identically, the field is called **irrotational**.

The curl operation is **distributive** but **not commutative** or **associative**, that is:

$$\nabla \times (\mathbf{A} + \mathbf{B}) = \nabla \times \mathbf{A} + \nabla \times \mathbf{B} \tag{37}$$

$$\nabla \times \mathbf{A} \neq \mathbf{A} \times \nabla \tag{38}$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) \neq (\nabla \times \mathbf{A}) \times \mathbf{B} \tag{39}$$

where **A** and **B** are differentiable vector functions of position.

The **Laplacian**[15] scalar operator ∇^2 is defined as the **divergence of the gradient** operator and hence it is given, in Cartesian coordinates, by:

$$\nabla^2 = \nabla \cdot \nabla = \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k} \right) \cdot \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k} \right) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{40}$$

The Laplacian can **act** on scalar, vector and tensor fields of higher rank. When the Laplacian operates on a tensor (in its general sense which includes scalar and vector) it produces a tensor of the **same rank**; hence the Laplacian of a scalar is a scalar, the Laplacian of a vector is a vector, the Laplacian of a rank-2 tensor is a rank-2 tensor, and so on.

F. Divergence Theorem

The divergence theorem, which is also known as **Gauss theorem**, is a mathematical statement of the intuitive idea that the **integral of the divergence** of a vector field over a given volume is **equal** to the **total flux** of the vector field out of the surface enclosing the volume. Symbolically, the divergence theorem states that:

$$\iiint_V \nabla \cdot \mathbf{A}\, d\tau = \iint_S \mathbf{A} \cdot \mathbf{n}\, d\sigma \tag{41}$$

[15] This operator is also known as the harmonic operator.

where **A** is a differentiable vector field, V is a bounded volume in an nD space enclosed by a surface S, $d\tau$ and $d\sigma$ are volume and surface elements respectively, and **n** is a variable unit vector normal to the surface.

The divergence theorem is useful for **converting volume integrals into surface integrals** and vice versa. In many cases, this can result in a considerable simplification of the required mathematical work when one of these integrals is easier to manipulate and evaluate than the other, or even overcoming a mathematical hurdle when one of the integrals cannot be evaluated analytically. Moreover, the divergence theorem plays a crucial role in many mathematical proofs and theoretical arguments in mathematical and physical theories.

G. Stokes Theorem

Stokes theorem is a mathematical statement that the **integral of the curl** of a vector field over an open surface is **equal** to the **line integral** of the field around the perimeter surrounding the surface, that is:

$$\iint_S (\nabla \times \mathbf{A}) \cdot \mathbf{n} \, d\sigma = \int_C \mathbf{A} \cdot d\mathbf{r} \tag{42}$$

where **A** is a differentiable vector field, C symbolizes the perimeter of the surface S, $d\mathbf{r}$ is a vector element tangent to the perimeter, and the other symbols are as defined in the divergence theorem. The perimeter should be traversed in a sense related to the direction of the normal vector **n** by the **right hand twist rule**, that is when the fingers of the right hand twist in the sense of traversing the perimeter the thumb will point approximately in the direction of **n**, as seen in Figure 10.

Similar to the divergence theorem, Stokes theorem is useful for **converting surface integrals into line integrals** and vice versa, which is useful in many cases for reducing the amount of mathematical work or overcoming technical and mathematical difficulties.

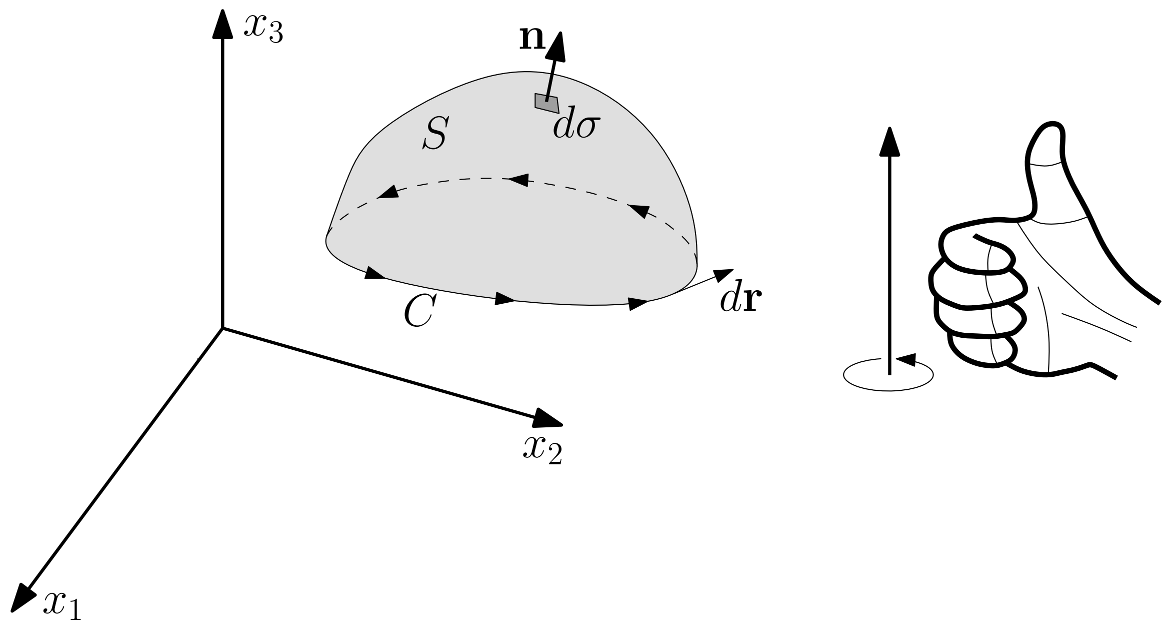


Figure 10: Illustration of Stokes integral theorem (left frame) with the right hand twist rule (right frame).

Stokes theorem is also crucial in the development of many proofs and theoretical arguments in mathematics and science.

1.3.3 Matrix Algebra

There is a close relation between rank-2 tensors and square matrices where the latter usually represent the former. Hence, there are many ideas, techniques and notations which are common or similar between the two subjects. We therefore provide in this subsection a set of short introductory notes about matrix algebra to supply the reader with the essential terminology and methods of matrix algebra which are needed in the development of the forthcoming chapters about tensor calculus.

A. Definition of Matrix

A matrix is a **rectangular array** of mathematical objects (mainly numbers or functions) which is subject to certain rules in its manipulation and over which certain mathematical operations are defined. Hence, **two indices** are needed to define a matrix unambiguously where the first index labels the **rows** while the second index labels the **columns**. A matrix which consists of m rows and n columns is said to be an $m \times n$ matrix.

The elements or entries of a matrix $\mathbf{A}$ are usually labeled with light-face symbols similar to the symbol used to label the matrix where each element is suffixed with two indices: the first refers to the row number of the entry and the second refers to its column number. Hence, for a matrix $\mathbf{A}$ the entry in its second row and fifth column is labeled A_{25}.

The two indices of a matrix are not required to have the same range since a matrix can have different number of rows and columns. When the two indices have the same range, the matrix is described as **square matrix**. Examples of matrices are:

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \qquad \begin{bmatrix} B_{11} \\ B_{21} \\ B_{31} \end{bmatrix} \qquad \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tag{43}$$

When the range of the row/column index has only one value (i.e. 1) while the range of the other index has multiple values the matrix is described as **row/column matrix**. Vectors may be represented by row or column matrices.[16] Scalars may be regarded as a trivial case of matrices.

For a square matrix, the entries with equal values of row and column indices are called the **main diagonal** of the matrix. For example, the entries of the main diagonal of the third matrix in Eq. 43 are C_{11} and C_{22}. The elements of the other diagonal running from the top right corner to the bottom left corner form the **trailing** or **anti-diagonal**, i.e. C_{12} and C_{21} in the previous example.

B. Special Matrices

[16] In this book, only square matrices are of primary interest as they are qualified to represent uniformly dimensioned rank-2 tensors. Row and column matrices are also qualified to represent employed vectors.

The **zero matrix** is a matrix whose all entries are 0. The **identity** or **unit** or **unity matrix** is a square matrix whose all entries are 0 except those on its main diagonal which are 1. A matrix is described as **singular** *iff* its determinant is zero (see § Determinant of Matrix). A singular matrix has no inverse (see § Inverse of Matrix). A square matrix is called **diagonal** if all of its elements which are not on the main diagonal are zero.

The **transpose** of a matrix is a matrix obtained by exchanging the rows and columns of the original matrix. For example, if $\mathbf{A}$ is a 3×3 matrix and $\mathbf{A}^T$ is its transpose then:[17]

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \qquad \Rightarrow \qquad \mathbf{A}^T = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \tag{44}$$

The transposition operation is defined even for non-square matrices. For square matrices, transposition represents a **reflection** of the matrix elements in the main diagonal of the matrix.

C. Matrix Multiplication

The multiplication of two matrices, $\mathbf{A}$ of $m \times k$ dimensions and $\mathbf{B}$ of $k \times n$ dimensions, is defined as an operation that produces a matrix $\mathbf{C}$ of $m \times n$ dimensions whose C_{ij} entry is the **dot product** of the i^{th} row of the first matrix $\mathbf{A}$ and the j^{th} column of the second matrix $\mathbf{B}$. Hence, if $\mathbf{A}$ is a 3×2 matrix and $\mathbf{B}$ is a 2×2 matrix, then their product $\mathbf{AB}$ is a 3×2 matrix which is given by:

$$\mathbf{AB} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \\ A_{31}B_{11} + A_{32}B_{21} & A_{31}B_{12} + A_{32}B_{22} \end{bmatrix} \tag{45}$$

[17] The indexing of the entries of $\mathbf{A}^T$ is not standard; the purpose of this is to demonstrate the exchange of rows and columns.

From the above, it can be seen that matrix multiplication is defined only when the number of columns of the first matrix is equal to the number of rows of the second matrix. Matrix multiplication is **associative** and **distributive** over a sum of compatible matrices, but it is **not commutative** in general even if both forms of the product are defined, that is:

$$(\mathbf{AB})\,\mathbf{C} = \mathbf{A}\,(\mathbf{BC}) \tag{46}$$

$$\mathbf{A}\,(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC} \tag{47}$$

$$\mathbf{AB} \neq \mathbf{BA} \tag{48}$$

As seen above, no symbol is used to indicate the operation of matrix multiplication according to the **notation of matrix algebra**, i.e. the two matrices are put side by side with no symbol in between. However, in tensor symbolic notation such an operation is usually represented by a dot between the symbols of the two matrices, as will be discussed later in the book.[18]

D. Trace of Matrix

The trace of a matrix is the **sum of its diagonal elements**, therefore if a matrix **A** is given by:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{49}$$

then its trace is given by:

$$\mathrm{tr}\,(\mathbf{A}) = A_{11} + A_{22} + A_{33} \tag{50}$$

[18] In brief, **AB** represents an inner product of **A** and **B** according to the matrix notation, and an outer product of **A** and **B** according to the symbolic notation of tensors, while $\mathbf{A}\cdot\mathbf{B}$ represents an inner product of **A** and **B** according to the symbolic notation of tensors. Hence, **AB** in the matrix notation is equivalent to $\mathbf{A}\cdot\mathbf{B}$ in the symbolic notation of tensors.

From its definition, it is obvious that the trace of a matrix is a **scalar** and it is **defined only for square matrices**.

E. Determinant of Matrix

The determinant is a **scalar** quantity associated with a **square matrix**. There are several definitions for the determinant of a matrix; the most direct one is that the determinant of a 2×2 matrix is the product of the elements of its main diagonal minus the product of the elements of its trailing diagonal, that is:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \Rightarrow \qquad \det(\mathbf{A}) = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}A_{22} - A_{12}A_{21} \tag{51}$$

The determinant of an $n \times n$ ($n > 2$) matrix is then defined, recursively, as the sum of the products of each entry of any one of its rows or columns times the **cofactor** of that entry where the cofactor of an entry is defined as the determinant obtained from eliminating the row and column of that entry from the parent matrix with a sign given by $(-1)^{i+j}$ with i and j being the indices of the row and column of that entry.[19] For example:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \\ &\Downarrow \\ \det(\mathbf{A}) &= A_{11} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{12} \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} + A_{13} \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} \end{aligned} \tag{52}$$

where the determinant is evaluated along the first row. It should be remarked that the

[19] In fact, this rule for the determinant of an $n \times n$ matrix applies even to the 2×2 matrix if the cofactor of an entry in this case is taken as a single entry with the designated sign. However, we separated the 2×2 matrix case in the definition to be more clear and to avoid possible confusion.

determinant of a matrix and the determinant of its transpose are equal, that is:

$$\det(\mathbf{A}) = \det\left(\mathbf{A}^T\right) \tag{53}$$

where $\mathbf{A}$ is a square matrix and T stands for the transposition operation. Another remark is that the determinant of a diagonal matrix is the product of its main diagonal elements.

F. Inverse of Matrix

The inverse of a **square matrix A** is a **square matrix $\mathbf{A}^{-1}$** where:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{54}$$

with $\mathbf{I}$ being the **identity matrix** (see § Special Matrices) of the same dimensions as $\mathbf{A}$. The inverse of a square matrix is formed by transposing the matrix of cofactors of the original matrix with dividing each element of the transposed matrix of cofactors by the determinant of the original matrix.[20] From this definition, it is obvious that a matrix possesses an inverse only if its determinant is not zero, i.e. it must be **non-singular**.

It should be remarked that this definition includes the 2×2 matrices where the cofactor of an entry is a single entry with the designated sign, that is:

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad \Rightarrow \qquad \mathbf{A}^{-1} = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix} \tag{55}$$

Another remark is that the inverse of an invertible diagonal matrix is a diagonal matrix obtained by taking the reciprocal of the corresponding diagonal elements of the original matrix.

[20] The matrix of cofactors (or cofactor matrix) is made of the cofactors of its elements taking the same positions as the positions of these elements. The transposed matrix of cofactors may be called the adjugate or adjoint matrix although the terminology may differ between the authors.

1.4 Exercises

1.1 Name three mathematicians accredited for the development of tensor calculus. For each one of these mathematicians, give a mathematical technical term that bears his name.

1.2 What are the main scientific disciplines that employ the language and techniques of tensor calculus?

1.3 Mention one cause for the widespread use of tensor calculus in science.

1.4 Describe some of the distinctive features of tensor calculus which contributed to its success and extensive use in mathematics, science and engineering.

1.5 Give preliminary definitions of the following terms: scalar, vector, tensor, rank of tensor, and dyad.

1.6 What is the meaning of the following mathematical symbols?

$$\partial_i \qquad \partial_{ii} \qquad \nabla \qquad A_{,i} \qquad \Delta \qquad A_{i;k}$$

1.7 Is the following equality correct? If so, is there any condition for this to hold?

$$\partial_k \partial_l = \partial_l \partial_k$$

1.8 Pick five terms from the Index about partial derivative operators and their symbols as used in tensor calculus and explain each one of these terms giving some examples for their use.

1.9 Describe, briefly, the following six coordinate systems outlining their main features:

orthonormal Cartesian, cylindrical, plane polar, spherical, general curvilinear, and general orthogonal.

1.10 Which of the six coordinate systems in the previous exercise are orthogonal?

1.11 What "basis vectors" of a coordinate system means and what purpose they serve?

1.12 Which of the six coordinate systems mentioned in the previous exercises have constant basis vectors (i.e. some or all of their basis vectors are constant both in magnitude and in direction)?

1.13 Which of the above six coordinate systems have unit basis vectors by definition or convention?

1.14 Explain the meaning of the coordinates in the cylindrical and spherical systems (i.e. ρ, ϕ and z for the cylindrical, and r, θ and ϕ for the spherical).

1.15 What is the relation between the cylindrical and plane polar coordinate systems?

1.16 Is there any common coordinates between the above six coordinate systems? If so, what? Investigate this thoroughly by comparing each pair of these systems.

1.17 Write the transformation equations between the following coordinate systems in both directions: Cartesian and cylindrical, and Cartesian and spherical.

1.18 Make a sketch representing a spherical coordinate system, with its basis vectors, superimposed on a rectangular Cartesian system in a standard position.

1.19 What are the geometric and algebraic definitions of the dot product of two vectors? What is the interpretation of the geometric definition?

1.20 What are the geometric and algebraic definitions of the cross product of two vectors? What is the interpretation of the geometric definition?

1.21 What is the dot product of the vectors $\mathbf{A}$ and $\mathbf{B}$ if $\mathbf{A} = (1.9, -6.3, 0)$ and $\mathbf{B} = (-4, -0.34, 11.9)$?

1.22 What is the cross product of the vectors in the previous exercise? Write this cross product in its determinantal form and expand it.

1.23 Define the scalar triple product operation of three vectors geometrically and algebraically.

1.24 What is the geometric interpretation of the scalar triple product? What is the condition for this product to be zero?

1.25 Is it necessary to use parentheses in the writing of scalar triple products and why? Is it possible to interchange the dot and cross symbols in the product?

1.26 Calculate the following scalar triple products:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \qquad \mathbf{a} \cdot (\mathbf{d} \times \mathbf{c}) \qquad \mathbf{d} \cdot (\mathbf{c} \times \mathbf{b}) \qquad \mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}) \qquad (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

where $\mathbf{a} = (7, -0.4, 9.5)$, $\mathbf{b} = (-12.9, -11.7, 3.1)$, $\mathbf{c} = (2.4, 22.7, -6.9)$ and $\mathbf{d} = (-56.4, 29.5, 33.8)$. Note that some of these products may be found directly from other products with no need for detailed calculations.

1.27 Write the twelve possibilities of the scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ and divide them into two sets where the entries in each set are equal. What is the relation between the two sets?

1.28 What is the vector triple product of three vectors in mathematical terms? Is it scalar or vector? Is it associative?

1.29 Give the mathematical expression for the nabla ∇ differential operator in Cartesian systems.

1.30 State the mathematical definition of the gradient of a scalar field f in Cartesian coordinates. Is it scalar or vector?

1.31 Is the gradient operation commutative, associative or distributive? Express these properties mathematically.

1.32 What is the relation between the gradient of a scalar field f and the surfaces of constant f? Make a simple sketch to illustrate this relation.

1.33 Define, mathematically, the divergence of a vector field $\mathbf{V}$ in Cartesian coordinates. Is it scalar or vector?

1.34 Define "solenoidal" vector field descriptively and mathematically.

1.35 What is the physical significance of the divergence of a vector field?

1.36 Is the divergence operation commutative, associative or distributive? Give your answer in words and in mathematical forms.

1.37 Define the curl of a vector field $\mathbf{V}$ in Cartesian coordinates using the determinantal and the expanded forms with full explanation of all the symbols involved. Is the curl scalar or vector?

1.38 What is the physical significance of the curl of a vector field?

1.39 What is the technical term used to describe a vector field whose curl vanishes identically?

1.40 Is the curl operation commutative, associative or distributive? Express these properties symbolically.

1.41 Describe, in words, the Laplacian operator ∇^2 and how it is obtained. What are the other symbols used to denote it?

1.42 Give the mathematical expression of the Laplacian operator in Cartesian systems. Using this mathematical expression, explain why the Laplacian is a scalar rather than a vector operator?

1.43 Can the Laplacian operator act on rank-0, rank-1 and rank-n ($n > 1$) tensor fields? If so, what is the rank of the resulting field in each case?

1.44 Collect from the Index all the terms related to the nabla based differential operations and classify them according to each one of these operations.

1.45 Write down the mathematical expression for the divergence theorem, defining all the symbols involved, and explain the meaning of this theorem in words.

1.46 What are the main uses of the divergence theorem in mathematics and science? Explain why this theorem is very useful theoretically and practically.

1.47 If a vector field is given in the Cartesian coordinates by $\mathbf{A} = (-0.5, 9.3, 6.5)$, verify the divergence theorem for a cube defined by the plane surfaces $x_1 = -1$, $x_2 = 1$, $y_1 = -1$, $y_2 = 1$, $z_1 = -1$, and $z_2 = 1$.

1.48 Write down the mathematical expression for Stokes theorem with the definition of all the symbols involved and explain its meaning in words. What is this theorem useful for? Why it is very useful?

1.49 If a vector field is given in the Cartesian coordinates by $\mathbf{A} = (2y, -3x, 1.5z)$, verify Stokes theorem for a hemispherical surface $x^2 + y^2 + z^2 = 9$ for $z \geq 0$.

1.50 Make a simple sketch to demonstrate Stokes theorem with sufficient explanations and definitions of the symbols involved.

1.51 Give concise definitions for the following terms related to matrices: matrix, square matrix, main diagonal, trailing diagonal, transpose, identity matrix, unit matrix, singular, trace, determinant, cofactor, and inverse.

1.52 Explain the way by which matrices are indexed.

1.53 How many indices are needed in indexing a 2×3 matrix, an $n \times n$ matrix, and an $m \times k$ matrix? Explain, in each case, why.

1.54 Does the order of the matrix indices matter? If so, what is the meaning of changing this order?

1.55 Is it possible to write a vector as a matrix? If so, what is the condition that should be imposed on the indices and how many forms a vector can have when it is written as a matrix?

1.56 Write down the following matrices in a standard rectangular array form (similar to the examples in Eq. 43) using conventional symbols for their entries with a proper indexing: 3×4 matrix $\mathbf{A}$, 1×5 matrix $\mathbf{B}$, 2×2 matrix $\mathbf{C}$, and 3×1 matrix $\mathbf{D}$.

1.57 Give detailed mathematical definitions of the determinant, trace and inverse of matrix, explaining any symbol or technical term involved in these definitions.

1.58 Find the following matrix multiplications: $\mathbf{AB}$, $\mathbf{BC}$, and $\mathbf{CB}$ where:

$$\mathbf{A} = \begin{bmatrix} 9.6 & 6.3 & -22 \\ -3.8 & 2.5 & 2.9 \\ -6 & 3.2 & 7.5 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} -3.8 & -2.0 \\ 4.6 & 11.6 \\ 12.0 & 25.9 \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} 3 & 8.4 & 61.3 \\ -5 & -33 & 5.9 \end{bmatrix}$$

1.59 Referring to the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ in the previous exercise, find all the permutations (repetitive and non-repetitive) involving two of these three matrices, and classify them into two groups: those which do represent possible matrix multiplication and those which do not.

1.60 Is matrix multiplication associative? commutative? distributive over matrix addition?

1.61 Calculate the trace, the determinant, and the inverse (if the inverse exists) of the following matrices:

$$\mathbf{D} = \begin{bmatrix} 3.2 & 2.6 & 1.6 \\ 12.9 & -1.9 & 2.4 \\ -11.9 & 33.2 & -22.5 \end{bmatrix} \qquad \mathbf{E} = \begin{bmatrix} 5.2 & 2.7 & 3.6 \\ -10.4 & -5.4 & -7.2 \\ -31.9 & 13.2 & -23.7 \end{bmatrix}$$

1.62 Which, if any, of the matrices **D** and **E** in the previous exercise is singular?

1.63 Select from the Index six terms connected to the special matrices which are defined in § Special Matrices.

Chapter 2

Tensors

In this chapter, we present the essential terms and definitions related to tensors, the conventions and notations which are used in their representation, the general rules that govern their manipulation, and their main types and classifications. We also provide some illuminating examples of tensors of various complexity as well as an overview of their use in mathematics, science and engineering.

2.1 General Background about Tensors

A tensor is an **array of mathematical objects** (usually numbers or functions) which transforms according to certain rules under coordinates change. In a d-dimensional space, a tensor of rank-n has d^n components which may be specified with reference to a given coordinate system. Accordingly, a scalar, such as temperature, is a **rank-0** tensor with (assuming a 3D space) $3^0 = 1$ component, a vector, such as force, is a **rank-1** tensor with $3^1 = 3$ components, and stress is a **rank-2** tensor with $3^2 = 9$ components. Figure 11 graphically illustrates the structure of a rank-3 tensor in a 3D space.

The d^n **components** of a tensor are identified by n distinct integer **indices** (e.g. i, j, k) which are attached, according to the commonly-employed **tensor notation**, as superscripts or subscripts or a mix of these to the right side of the symbol utilized to label the tensor, e.g. A_{ijk}, A^{ijk} and A_i^{jk}. Each tensor index takes all the values over a predefined **range** of dimensions such as 1 to d in the above example of a d-dimensional space. In gen-

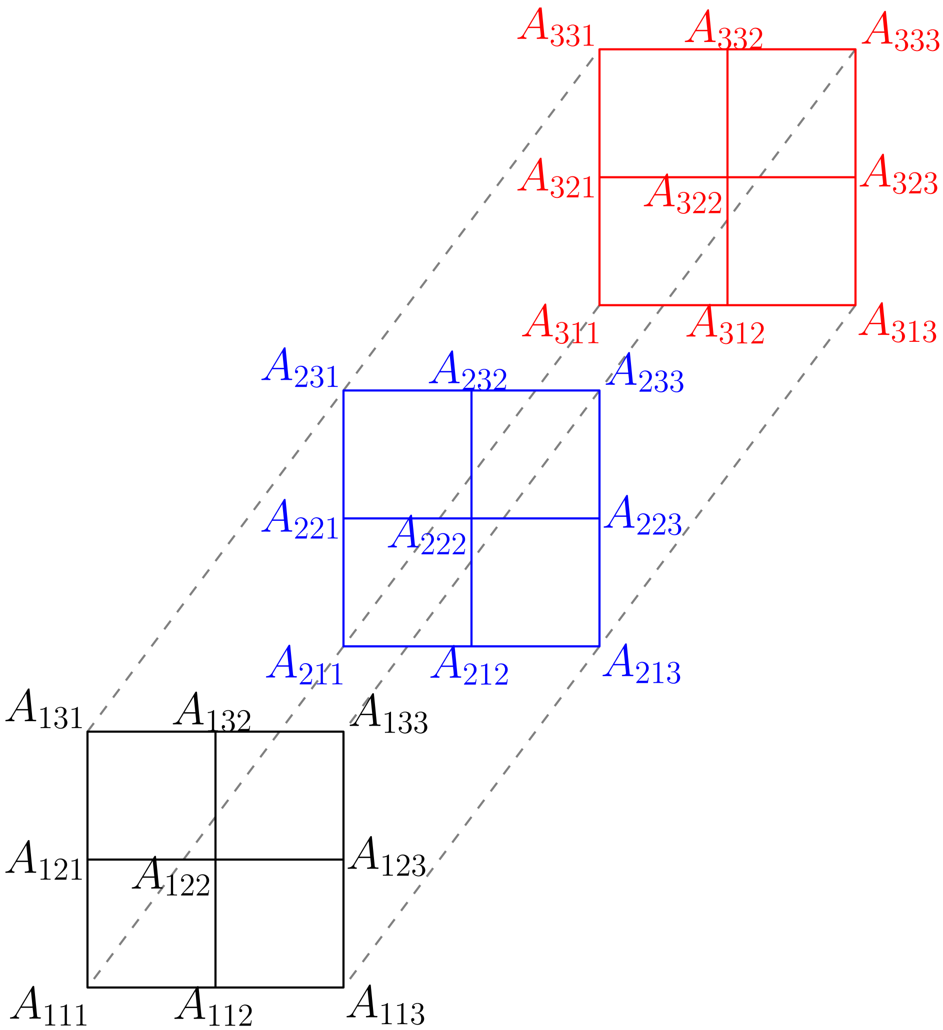


Figure 11: Graphical illustration of a rank-3 tensor A_{ijk} in a 3D space, i.e. each one of i, j, k ranges over $1, 2, 3$.

eral, all tensor indices have the same range, i.e. they are uniformly dimensioned.[21] When the range of tensor indices is not stated explicitly, it is usually assumed to have the values $1, 2, 3$. However, the range must be stated explicitly or implicitly to avoid ambiguity.

The characteristic property of tensors is that they satisfy the **principle of invariance** under certain coordinate transformations. Therefore, formulating the fundamental physical laws in a tensor form ensures that they are **form-invariant**; hence they are

[21] This assertion, in fact, applies to the common cases of tensor applications; however, there are instances, for example in the differential geometry of curves and surfaces, of tensors which are not uniformly dimensioned because the tensor is related to two spaces with different dimensions.

objectively-representing the physical reality and do not depend on the observer. Having the same form in different coordinate systems may be labeled as being **covariant** although this word is also used for a different meaning in tensor calculus, as will be fully explained in § 2.6.1.

While tensors of rank-0 are generally represented in a common form of light face non-indexed italic symbols like f and g, tensors of rank ≥ 1 are represented in several forms and notations, the main ones are the **index-free notation**, which may also be called the **direct** or **symbolic** or **Gibbs notation**, and the **indicial notation** which is also called the **index** or **component** or **tensor notation**. The first is a **geometrically** oriented notation with no reference to a particular reference frame and hence it is intrinsically invariant to the choice of coordinate systems, whereas the second takes an **algebraic** form based on components identified by indices and hence the notation is suggestive of an underlying coordinate system, although being a tensor makes it form-invariant under certain coordinate transformations and therefore it possesses certain invariant properties. The index-free notation is usually identified by using **bold face** non-italic symbols, like **a** and **B**, while the indicial notation is identified by using **light face** indexed italic symbols such as a^i and B_{ij}.

It is noteworthy that although rank-0 and rank-1 tensors are, respectively, scalars and vectors, not all scalars and vectors (in their generic sense) are tensors of these ranks. Similarly, rank-2 tensors are normally represented by square matrices but not all square matrices represent rank-2 tensors.

2.2 General Terms and Concepts

In the following, we introduce and define a number of essential concepts and terms which form a principal part of the technical and conceptual structure of tensor calculus. These concepts and terms are needed in the development of the forthcoming sections and chapters.

Tensor **term** is a product of tensors including scalars and vectors. Tensor **expression** is an algebraic sum (or more generally a linear combination) of tensor terms which may be a trivial sum in the case of a single term. Tensor **equality** (which is symbolized by "=") is an equality of two tensor terms and/or expressions. A special case of this is tensor **identity** which is an equality of general validity.[22]

An index that occurs once in a tensor term is a **free index** while an index that occurs twice in a tensor term is a **dummy** or **bound index**. The **order** of a tensor is identified by the number of its indices (e.g. A^i_{jk} is a tensor of order 3) which normally identifies the tensor **rank** as well. However, when contraction (see § 3.4) of indices operation takes place once or more, the order of the tensor is not affected but its rank is reduced by two for each contraction operation.[23] Hence, the order of a tensor is equal to the number of all of its indices including the dummy indices, while the rank is equal to the number of its free indices only.

Tensors with subscript indices, like A_{ij}, are called **covariant**, while tensors with superscript indices, like A^k, are called **contravariant**. Tensors with both types of indices, like A^{lmn}_{lk}, are called **mixed** type. More details about this classification will follow in § 2.6.1. Subscript indices, rather than subscripted tensors, are also dubbed covariant and superscript indices are dubbed contravariant.

The **Zero** tensor is a tensor whose all components are zero. The **Unit** tensor or **unity** tensor, which is usually defined for rank-2 tensors, is a tensor whose all elements are zero except the ones with identical values of all indices which are assigned the value 1.

In general terms, a **transformation** from an nD space to another nD space is a cor-

[22] As indicated previously, the symbol "≡" may be used for identity as well as for definition.

[23] In the literature of tensor calculus, rank and order of tensors are generally used interchangeably; however some authors differentiate between the two as they assign order to the total number of indices, including contracted indices, while they reserve rank to the number of free indices. We think the latter is better and hence in the present book we embrace this terminology.

relation that maps a point from the first space (original) to a point in the second space (transformed) where each point in the original and transformed spaces is identified by n **independent** variables or coordinates. To distinguish between the two sets of coordinates in the two spaces, the coordinates of the points in the **transformed space** may be notated with **barred** symbols like $(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^n)$ or $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ where the superscripts and subscripts are indices, while the coordinates of the points in the **original space** are notated with **unbarred** symbols like $(x^1, x^2, \ldots, x^n)$ or $(x_1, x_2, \ldots, x_n)$. Under certain conditions, such a transformation is unique and hence an **inverse transformation** from the transformed to the original space is also defined.

Mathematically, each one of the **direct** and **inverse transformations** can be regarded as a mathematical correlation expressed by a set of equations in which each coordinate in one space is considered as a **function** of the coordinates in the other space. Hence, the transformations between the two sets of coordinates in the two spaces can be expressed mathematically by the following two sets of **independent** relations:

$$\bar{x}^i = \bar{x}^i(x^1, x^2, \ldots, x^n) \tag{56}$$

$$x^i = x^i(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^n) \tag{57}$$

where $i = 1, 2, \ldots, n$.

An **alternative** to the latter view of considering the transformation as a mapping between two different spaces is to view it as a correlation relating the same point in the same space but observed from two different coordinate systems which are subject to a similar transformation. The following will be largely based on the latter view.

Coordinate transformations are described as **proper** when they preserve the **handedness** (right- or left-handed) of the coordinate system and **improper** when they reverse the handedness. Improper transformations involve an odd number of coordinate axes inversions in the origin of coordinates. Inversion of axes may be called **improper rotation** while ordinary rotation is described as **proper rotation**. Figure 12 illustrates proper and improper coordinate transformations of a rectangular Cartesian system.

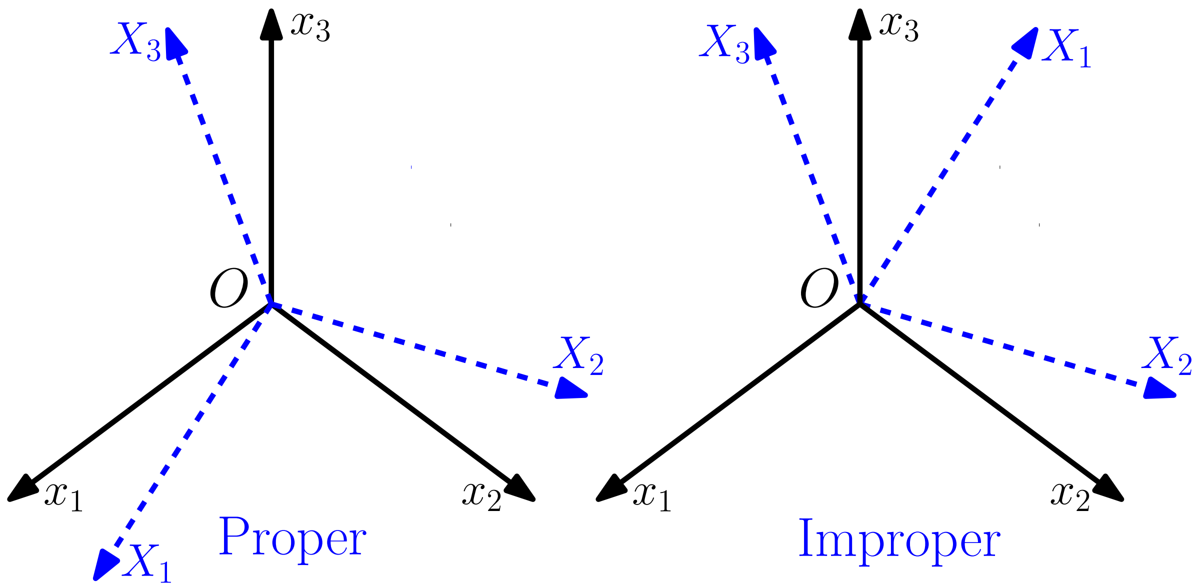


Figure 12: Proper and improper transformations of a rectangular Cartesian coordinate system where the former is achieved by a rotation of the coordinate system while the latter is achieved by a rotation followed by a reflection of the first axis in the origin of coordinates. The transformed systems are shown as dashed while the original system is shown as solid.

Transformations can be **active**, when they change the state of the observed object (e.g. translating the object in space), or **passive** when they are based on keeping the state of the object and changing the state of the coordinate system from which the object is observed. Such distinction is based on an implicit assumption of a more general frame of reference in the background.

A **permutation** of a set of objects, which are normally numbers like $(1, 2, \ldots, n)$ or symbols like (i, j, k), is a particular ordering or arrangement of these objects. An **even** permutation is a permutation resulting from an even number of single-step **exchanges** (also known as **transpositions**) of neighboring objects starting from a presumed original permutation of these objects. Similarly, an **odd** permutation is a permutation resulting from an odd number of such exchanges. It has been shown that when a transformation from one permutation to another can be done in different ways, possibly with different numbers of exchanges, the **parity** of all these possible transformations is the same, i.e. all are even or all are odd, and hence there is no ambiguity in characterizing the transformation from one permutation to another by the parity alone.

2.3 General Rules

In the following, we present some very general rules that apply to the mathematical expressions and relations in tensor calculus. No index is allowed to occur **more than twice** in a legitimate tensor term.[24] A free index should be understood to **vary over its range** (e.g. $1, \ldots, n$) and hence it can be interpreted as saying "for all components represented by the index". Therefore, a free index represents a number of terms or expressions or equalities equal to the number of allowed values of its range. For example, when i and j can vary over the range $1, \ldots, n$ the following expression:

$$A_i + B_i \tag{58}$$

represents n separate expressions while the following equation:

$$A_i^j = B_i^j \tag{59}$$

represents $n \times n$ separate equations.

[24] We adopt this assertion, which is common in the literature of tensor calculus, as we think it is suitable for this level. However, there are many instances in the literature of tensor calculus where indices are repeated more than twice in a single term. The bottom line is that as long as the tensor expression makes sense and the intention is clear, such repetitions should be allowed with no need in our view to take special precaution like using parentheses. In particular, the forthcoming summation convention will not apply automatically in such cases although summation on such indices, if needed, can be carried out explicitly, by using the summation symbol $\sum$ or by a special declaration of such intention similar to the summation convention. Anyway, in the present book we will not use indices repeated more than twice in a single term.

According to the **summation convention**, which is widely used in the literature of tensor calculus including in the present book, **dummy indices** imply summation over their range, e.g. for an nD space we have:

$$A^i B_i = \sum_{i=1}^{n} A^i B_i = A^1 B_1 + A^2 B_2 + \ldots + A^n B_n \tag{60}$$

$$\delta_{ij} A^{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij} A^{ij} \tag{61}$$

$$\epsilon_{ijk} A^{ij} B^k = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \epsilon_{ijk} A^{ij} B^k \tag{62}$$

When dummy indices do not imply summation, the situation must be clarified by enclosing such indices in **parentheses** or by **underscoring** or by using **upper case** letters (with declaration of these conventions) or by adding a clarifying **comment** like "no summation on repeated indices".[25]

Each tensor index should conform to one of the forthcoming variance transformation rules as given by Eqs. 70 and 71, i.e. it is either **covariant** or **contravariant**. For orthonormal **Cartesian** coordinate systems, the two variance types (i.e. covariant and contravariant) **do not differ** because the metric tensor is given by the Kronecker delta (refer to § 4.1 and 6) and hence any index can be upper or lower although it is common to use lower indices in such cases.

For tensor invariance, a pair of dummy indices should in general be **complementary** in their variance type, i.e. one covariant and the other contravariant. However, for orthonormal Cartesian systems the two are the same and hence when both dummy indices are covariant or both are contravariant it should be understood as an indication that the underlying coordinate system is orthonormal Cartesian if the possibility of an error is

[25] These precautions are obviously needed if the summation convention is adopted in general but it does not apply in some exceptional cases where repeated indices are needed in the notation with no intention of summation.

excluded.

As indicated earlier, tensor **order** is equal to the number of its indices while tensor **rank** is equal to the number of its free indices; hence vectors (terms, expressions and equalities) are represented by a single free index and rank-2 tensors are represented by two free indices. The **dimension** of a tensor is determined by the range taken by its indices.

The **rank** of all terms in legitimate tensor expressions and equalities must be the same. Moreover, each term in valid tensor expressions and equalities must have the same set of free indices (e.g. i, j, k). Also, a **free index** should **keep** its variance type in every term in valid tensor expressions and equations, i.e. it must be covariant in all terms or contravariant in all terms.

While free indices should be named **uniformly** in all terms of tensor expressions and equalities, dummy indices can be named in each term **independently**, e.g.

$$A^i_{ik} + B^j_{jk} + C^{lm}_{lmk} \qquad\qquad D^j_i = E^{jk}_{ik} + F^{jm}_{im} \tag{63}$$

A free index in an expression or equality can be renamed uniformly using a different symbol, as long as this symbol is not already in use, assuming that both symbols vary over the same range, i.e. have the same dimension.

Examples of **legitimate** tensor terms, expressions and equalities are:

$$A^{ij}_{ij} \qquad A^{im}_m + B^{ink}_{nk} \qquad C_{ij} = A_{ij} - B_{ij} \qquad a = B^j_j \tag{64}$$

while examples of **illegitimate** tensor terms, expressions and equalities are:

$$B^{ii}_i \qquad A_i + B_{ij} \qquad A^i + B^j \qquad A_i - B^i \qquad A^i_i = B_i \tag{65}$$

Indexing is generally **distributive** over the terms of tensor expressions and equalities, for example:

$$[\mathbf{A} + \mathbf{B}]_i = [\mathbf{A}]_i + [\mathbf{B}]_i \tag{66}$$

and

$$[\mathbf{A} = \mathbf{B}]_i \qquad \Longleftrightarrow \qquad [\mathbf{A}]_i = [\mathbf{B}]_i \tag{67}$$

Unlike scalars and tensor components, which are essentially scalars in a generic sense, **operators** cannot in general be freely reordered in tensor terms, therefore we have:

$$fh = hf \qquad\qquad A_i B^i = B^i A_i \tag{68}$$

but:

$$\partial_i A_i \neq A_i \partial_i \tag{69}$$

It should be remarked that the **order of the indices**[26] of a given tensor is important and hence it should be observed and clarified, because two tensors with the same set of indices and with the same indicial structure but with different indicial order are not equal in general. For example, A_{ijk} is not equal to A_{jik} unless $\mathbf{A}$ is symmetric (refer to § 2.6.5) with respect to the indices i and j. Similarly, B^{mln} is not equal to B^{lmn} unless $\mathbf{B}$ is symmetric in its indices l and m.

The confusion about the order of indices occurs, in particular, in the case of mixed type tensors such as A^i_{jk}. Spaces are usually used in this case to indicate the order, e.g. the latter tensor is symbolized as $A_j{}^i{}_k$ if the order of the indices is j, i, k while it is symbolized as $A^i{}_{jk}$ if the order of the indices is i, j, k.[27] **Dots** may also be used in the case of mixed type tensors to indicate, more explicitly, the order of the indices and remove

[26] This should not be confused with the order of tensor as defined above in the same context as tensor rank.
[27] We exaggerate the spacing here for clarity.

any ambiguity. For example, if the indices i, j, k of the tensor **A**, which is covariant in i and k and contravariant in j, are of that order, then **A** may be symbolized as $A_{i\ \cdot\ k}^{\ j}$ where the dot between i and k indicates that j is in the middle.[28]

Finally, many of the identities in the present book which are given in a covariant or a contravariant or a mixed form are **similarly valid** for the other forms and hence they can be obtained with a minimal effort. The objective of reporting in only one form is conciseness and to avoid unnecessary repetition. Moreover, in the case of orthonormal Cartesian systems the variance type of indices is irrelevant.

2.4 Examples of Tensors of Different Ranks

Examples of **rank-0** tensors (scalars) are energy, mass, temperature, volume and density. These are totally identified by a single number regardless of any coordinate system and hence they are invariant under coordinate transformations.[29]

Examples of **rank-1** tensors (vectors) are displacement, force, electric field, velocity and acceleration. These require for their complete identification a number, representing their magnitude, and a direction representing their geometric orientation within their space. Alternatively, they can be uniquely identified by a set of numbers, equal to the number of dimensions of the underlying space, in reference to a particular coordinate system and hence this identification is system-dependent although they still have system-invariant properties such as length.

Examples of **rank-2** tensors are the Kronecker delta (see § 4.1), stress, strain, rate of

[28] In many places in this book (like other books) and for the convenience in typesetting, the order of the indices is not clarified by spacing or inserting dots in the case of mixed type tensors. This commonly occurs where the order of the indices is irrelevant in the given context or the order is clear. Sometimes, the order of the indices may be indicated implicitly by the alphabetical order of the selected indices, e.g. A_i^{jk} means $A_i^{\ jk}$.

[29] The focus of this section is on providing examples of tensors of different ranks. As we will see later in this chapter (refer to § 2.6.2), there are true and pseudo scalars, vectors and tensors, and hence some of the statements and examples given here may qualify for certain restrictions and conditions.

strain and inertia tensors. These require for their full identification a set of numbers each of which is associated with two directions. These **double directions** are usually identified by a set of **unit dyads**. Figure 13 is a graphic illustration of the nine unit dyads which are associated with the double directions of rank-2 tensors in a 3D space with a rectangular Cartesian coordinate system.

Examples of **rank-3** tensors are the Levi-Civita tensor (see § 4.2) in 3D spaces and the tensor of piezoelectric moduli. Examples of **rank-4** tensors are the elasticity or stiffness tensor, the compliance tensor and the fourth-order moment of inertia tensor. It is noteworthy that tensors of high ranks are relatively rare in science and engineering.

2.5 Applications of Tensors

Tensor calculus is a very powerful mathematical tool; hence tensors are commonplace in science and engineering where they are used to represent physical and synthetic objects and ideas in mathematical invariant forms. Tensor notation and techniques are used in many branches of mathematics, science and engineering such as **differential geometry**, **fluid mechanics**, **continuum mechanics**, **general relativity** and **structural engineering**. Tensor calculus is used for elegant and compact formulation and presentation of equations and identities in mathematics, science and engineering. It is also used for algebraic manipulation of mathematical expressions and proving identities in a neat and succinct way (refer to § 5.6). As indicated earlier, the invariance of tensor forms serves a theoretically and practically important role by allowing the formulation of physical laws in coordinate-free forms.

2.6 Types of Tensors

In the following subsections we introduce a number of tensor types and categories and highlight their main characteristics and differences. These types and categories are **not**

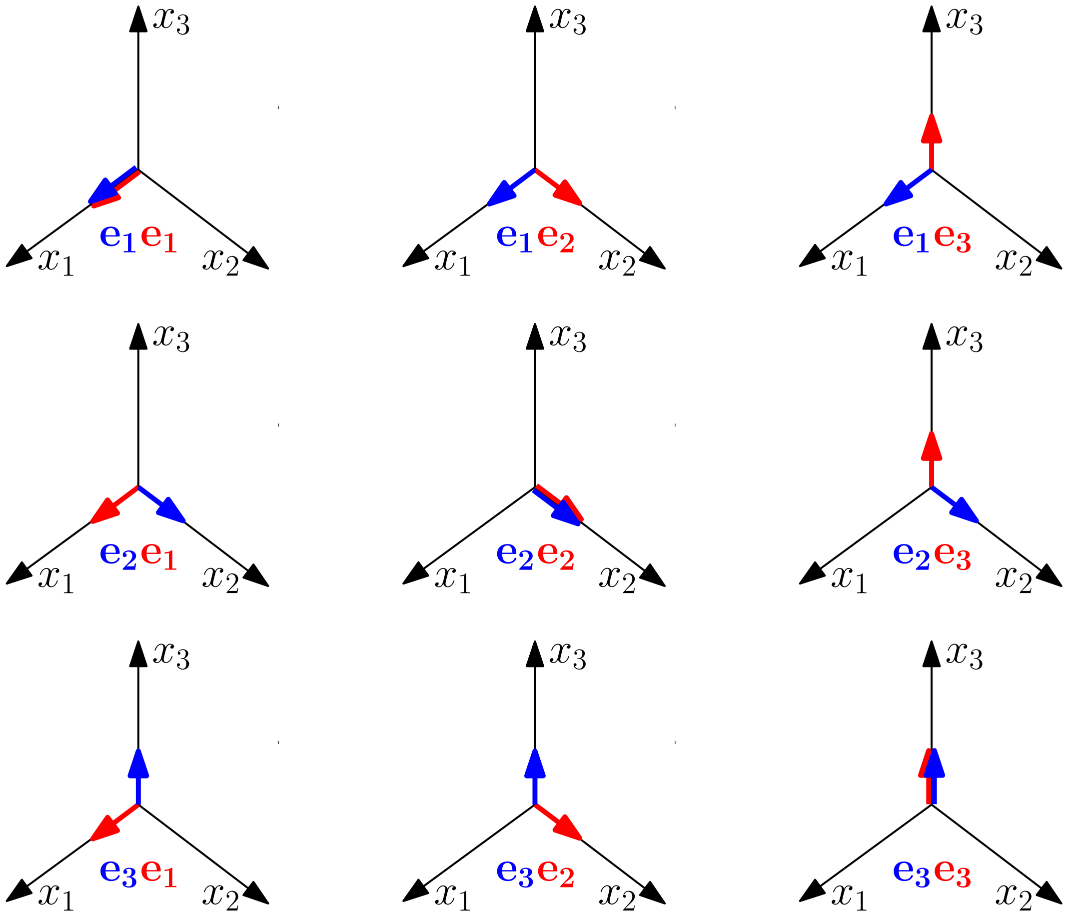


Figure 13: The nine unit dyads associated with the double directions of rank-2 tensors in a 3D space with a rectangular Cartesian coordinate system. The vectors $\mathbf{e}_i$ and $\mathbf{e}_j$ $(i, j = 1, 2, 3)$ are unit vectors in the directions of coordinate axes where the first indexed $\mathbf{e}$ (blue) represents the first vector of the dyad while the second indexed $\mathbf{e}$ (red) represents the second vector of the dyad. In these nine frames, the first vector is fixed along each row while the second vector is fixed along each column.

mutually exclusive and hence they overlap in general; moreover they may not be exhaustive in their classes as some tensors may not instantiate any one of a complementary set of types such as being symmetric or anti-symmetric.

2.6.1 Covariant and Contravariant Tensors

These are the main types of tensor with regard to the rules of their transformation between different coordinate systems. **Covariant** tensors are notated with **subscript** indices (e.g.

A_i) while **contravariant** tensors are notated with **superscript** indices (e.g. A^{ij}). A covariant tensor is transformed according to the following rule:

$$\bar{A}_i = \frac{\partial x^j}{\partial \bar{x}^i} A_j \tag{70}$$

while a contravariant tensor is transformed according to the following rule:

$$\bar{A}^i = \frac{\partial \bar{x}^i}{\partial x^j} A^j \tag{71}$$

where the barred and unbarred symbols represent the same mathematical object (tensor or coordinate) in the transformed and original coordinate systems respectively.

An example of covariant tensors is the **gradient** of a scalar field while an example of contravariant tensors is the **displacement** vector. Some tensors of rank > 1 have **mixed** variance type, i.e. they are covariant in some indices and contravariant in others. In this case the covariant variables are indexed with subscripts while the contravariant variables are indexed with superscripts, e.g. A_i^j which is covariant in i and contravariant in j. A mixed type tensor transforms covariantly in its covariant indices and contravariantly in its contravariant indices, e.g.

$$\bar{A}^{l\ \ n}_{\ m} = \frac{\partial \bar{x}^l}{\partial x^i} \frac{\partial x^j}{\partial \bar{x}^m} \frac{\partial \bar{x}^n}{\partial x^k} A^{i\ \ k}_{\ j} \tag{72}$$

To clarify the **pattern of mathematical transformation** of tensors, we explain step-by-step the practical rules to follow in writing tensor transformation equations between two coordinate systems, unbarred and barred, where for clarity we color the symbols of the tensor and the coordinates belonging to the unbarred system with blue while we use red to mark the symbols belonging to the barred system. Since there are three types of tensors: covariant, contravariant and mixed, we use three equations in each step.

In this demonstration we use rank-4 tensors as examples since this is sufficiently general and hence adequate to elucidate the rules for transforming tensors of any rank. The demonstration is based on the assumption that the **transformation** is taking place from the unbarred system to the barred system; the same rules should apply for the opposite transformation from the barred system to the unbarred system. We use the sign "$\circeq$" for the equality in the transitional steps to indicate that the equalities are under construction and are not complete.

We start with the very generic equations between the barred tensor $\bar{A}$ and the unbarred tensor A for the three types:

$$\begin{aligned} \bar{A} &\circeq A \qquad \text{(covariant)} \\ \bar{A} &\circeq A \qquad \text{(contravariant)} \\ \bar{A} &\circeq A \qquad \text{(mixed)} \end{aligned} \tag{73}$$

We assume that the barred tensor and its coordinates are indexed with $ijkl$ and the unbarred are indexed with $npqr$, so we add these indices in their presumed order and position (lower or upper) paying particular attention to the order in the mixed type:

$$\begin{aligned} \bar{A}_{ijkl} &\circeq A_{npqr} \\ \bar{A}^{ijkl} &\circeq A^{npqr} \\ \bar{A}^{ij}{}_{kl} &\circeq A^{np}{}_{qr} \end{aligned} \tag{74}$$

Since the barred and unbarred tensors are of the same type, as they represent the **same tensor** in two coordinate systems,[30] the indices on the two sides of the equalities should **match** in their position and order. We then insert a number of partial differential operators on the right hand side of the equations equal to the rank of these tensors, which is 4 in our

[30] Similar basis vectors are assumed.

example. These operators represent the transformation rules for each pair of corresponding coordinates, one from the barred and one from the unbarred:

$$\begin{aligned} \bar{A}_{ijkl} &\doteq \frac{\partial}{\partial}\frac{\partial}{\partial}\frac{\partial}{\partial}\frac{\partial}{\partial} \quad A_{npqr} \\ \bar{A}^{ijkl} &\doteq \frac{\partial}{\partial}\frac{\partial}{\partial}\frac{\partial}{\partial}\frac{\partial}{\partial} \quad A^{npqr} \\ \bar{A}^{ij}{}_{kl} &\doteq \frac{\partial}{\partial}\frac{\partial}{\partial}\frac{\partial}{\partial}\frac{\partial}{\partial} \quad A^{np}{}_{qr} \end{aligned} \tag{75}$$

Now we insert the coordinates of the barred system into the partial differential operators noting that (i) the positions of any index on the two sides should match, i.e. both upper or both lower, since they are free indices in different terms of tensor equalities, (ii) a **superscript** index in the **denominator** of a partial derivative is in lieu of a **covariant** index in the **numerator**,[31] and (iii) the order of the coordinates should match the order of the indices in the tensor, that is:

$$\begin{aligned} \bar{A}_{ijkl} &\doteq \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}\frac{\partial}{\partial x^k}\frac{\partial}{\partial x^l} \quad A_{npqr} \\ \bar{A}^{ijkl} &\doteq \frac{\partial x^i}{\partial}\frac{\partial x^j}{\partial}\frac{\partial x^k}{\partial}\frac{\partial x^l}{\partial} \quad A^{npqr} \\ \bar{A}^{ij}{}_{kl} &\doteq \frac{\partial x^i}{\partial}\frac{\partial x^j}{\partial}\frac{\partial}{\partial x^k}\frac{\partial}{\partial x^l} \quad A^{np}{}_{qr} \end{aligned} \tag{76}$$

For consistency, these coordinates should be barred as they belong to the barred tensor; hence we add bars:

$$\begin{aligned} \bar{A}_{ijkl} &\doteq \frac{\partial}{\partial \bar{x}^i}\frac{\partial}{\partial \bar{x}^j}\frac{\partial}{\partial \bar{x}^k}\frac{\partial}{\partial \bar{x}^l} \quad A_{npqr} \\ \bar{A}^{ijkl} &\doteq \frac{\partial \bar{x}^i}{\partial}\frac{\partial \bar{x}^j}{\partial}\frac{\partial \bar{x}^k}{\partial}\frac{\partial \bar{x}^l}{\partial} \quad A^{npqr} \\ \bar{A}^{ij}{}_{kl} &\doteq \frac{\partial \bar{x}^i}{\partial}\frac{\partial \bar{x}^j}{\partial}\frac{\partial}{\partial \bar{x}^k}\frac{\partial}{\partial \bar{x}^l} \quad A^{np}{}_{qr} \end{aligned} \tag{77}$$

[31] The use of upper indices in the symbols of general coordinates is to indicate the fact that the coordinates and their differentials transform contravariantly.

Finally, we insert the coordinates of the unbarred system into the partial differential operators noting that (i) the **positions** of the repeated indices on the same side should be **opposite**, i.e. one upper and one lower, since they are dummy indices and hence the position of the index of the unbarred coordinate should be opposite to its position in the unbarred tensor, (ii) an upper index in the denominator is in lieu of a lower index in the numerator, and (iii) the order of the coordinates should **match** the order of the indices in the tensor:

$$\begin{aligned}
\bar{A}_{ijkl} &= \frac{\partial x^n}{\partial \bar{x}^i}\frac{\partial x^p}{\partial \bar{x}^j}\frac{\partial x^q}{\partial \bar{x}^k}\frac{\partial x^r}{\partial \bar{x}^l} \quad A_{npqr} \\
\bar{A}^{ijkl} &= \frac{\partial \bar{x}^i}{\partial x^n}\frac{\partial \bar{x}^j}{\partial x^p}\frac{\partial \bar{x}^k}{\partial x^q}\frac{\partial \bar{x}^l}{\partial x^r} \quad A^{npqr} \\
\bar{A}^{ij}{}_{kl} &= \frac{\partial \bar{x}^i}{\partial x^n}\frac{\partial \bar{x}^j}{\partial x^p}\frac{\partial x^q}{\partial \bar{x}^k}\frac{\partial x^r}{\partial \bar{x}^l} \quad A^{np}{}_{qr}
\end{aligned} \tag{78}$$

We also replaced the "$\mathrel{\hat{=}}$" sign in the final set of equations with the strict equality sign "$=$" as the equations now are complete.

The covariant and contravariant types of a tensor are linked through the **metric tensor**, as will be detailed later in the book (refer to § 6). As indicated before, for orthonormal **Cartesian** systems there is **no difference** between covariant and contravariant tensors, and hence the indices can be upper or lower although it is common to use lower indices in this case.

A tensor of m contravariant indices and n covariant indices may be called **type** (m, n) tensor. When one or both variance types are absent, **zero** is used to refer to the absent variance type in this notation. Accordingly, A^k_{ij} is a type $(1, 2)$ tensor, B^{ik} is a type $(2, 0)$ tensor, C_m is a type $(0, 1)$ tensor, and D^{ts}_{pqr} is a type $(2, 3)$ tensor.

The vectors providing the **basis set** for a coordinate system are of **covariant** type when they are **tangent** to the coordinate axes, and they are of **contravariant** type when they are **perpendicular** to the local surfaces of constant coordinates. These two sets, like the tensors themselves, are identical for orthonormal Cartesian systems.

Formally, the covariant and contravariant basis vectors are given respectively by:

$$\mathbf{E}_i = \frac{\partial \mathbf{r}}{\partial x^i} \qquad\qquad \mathbf{E}^i = \nabla x^i \tag{79}$$

where $\mathbf{r} = x_i \mathbf{e}_i$ is the position vector in Cartesian coordinates and x^i is a general curvilinear coordinate. As before, a **superscript** in the **denominator** of partial derivatives is equivalent to a **subscript** in the **numerator**. It should be remarked that in general the basis vectors (whether covariant or contravariant) are **not necessarily** of **unit length** and/or **mutually orthogonal** although they may be so.[32]

The two sets of covariant and contravariant basis vectors are reciprocal systems and hence they satisfy the following **reciprocity relation**:

$$\mathbf{E}_i \cdot \mathbf{E}^j = \delta_i^j \tag{80}$$

where δ_i^j is the **Kronecker** delta (refer to § 4.1) which can be represented by the **unity** matrix (see § Special Matrices). The reciprocity of these two sets of basis vectors is illustrated schematically in Figure 14 for the case of a 2D space.

A vector can be represented either by **covariant components** with contravariant coordinate basis vectors or by **contravariant components** with covariant coordinate basis vectors. For example, a vector $\mathbf{A}$ can be expressed as:

$$\mathbf{A} = A_i \mathbf{E}^i \qquad \text{or} \qquad \mathbf{A} = A^i \mathbf{E}_i \tag{81}$$

where $\mathbf{E}^i$ and $\mathbf{E}_i$ are the contravariant and covariant basis vectors respectively. This is illustrated graphically in Figure 15 for a vector $\mathbf{A}$ in a 2D space. The use of the covariant

[32] In fact there are standard mathematical procedures to orthonormalize the basis set if it is not and if this is needed.

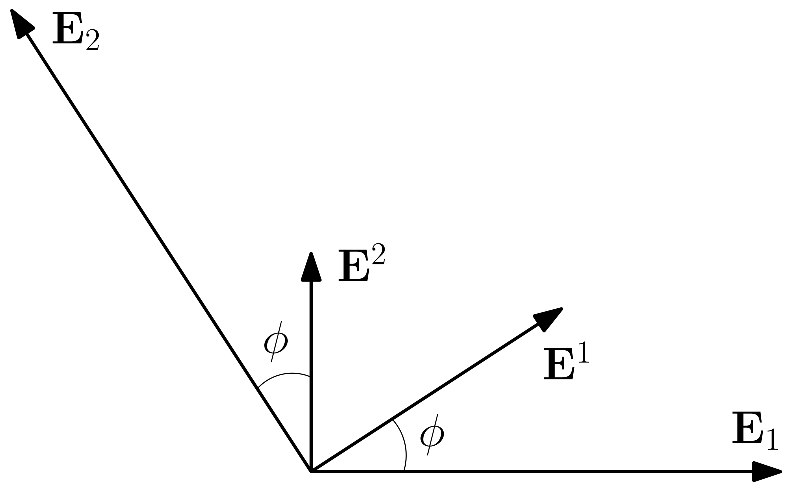


Figure 14: The reciprocity relation between the covariant and contravariant basis vectors in a 2D space where $\mathbf{E}_1 \perp \mathbf{E}^2$, $\mathbf{E}^1 \perp \mathbf{E}_2$, and $|\mathbf{E}_1|\,|\mathbf{E}^1|\cos\phi = |\mathbf{E}_2|\,|\mathbf{E}^2|\cos\phi = 1$.

or contravariant form of the vector representation is a **matter of choice** and convenience since these two representations are equivalent as they represent and correctly describe the same object.

More generally, a tensor of any rank (≥ 1) can be represented **covariantly** using contravariant basis tensors of that rank, or **contravariantly** using covariant basis tensors, or in a **mixed** form using a mixed basis of opposite type. For example, a rank-2 tensor $\mathbf{A}$ can be written as:

$$\mathbf{A} = A_{ij}\mathbf{E}^i\mathbf{E}^j = A^{ij}\mathbf{E}_i\mathbf{E}_j = A_i{}^j\mathbf{E}^i\mathbf{E}_j = A^i{}_j\mathbf{E}_i\mathbf{E}^j \tag{82}$$

where $\mathbf{E}^i\mathbf{E}^j$, $\mathbf{E}_i\mathbf{E}_j$, $\mathbf{E}^i\mathbf{E}_j$ and $\mathbf{E}_i\mathbf{E}^j$ are dyadic products of the basis vectors of the presumed system (refer to § 2.4 and 3.3).

2.6.2 True and Pseudo Tensors

These are also called **polar** and **axial** tensors respectively although it is more common to use these terms for vectors. Pseudo tensors may also be called **tensor densities**.[33] True

[33] The terminology in this part, like many other parts, is not universal.

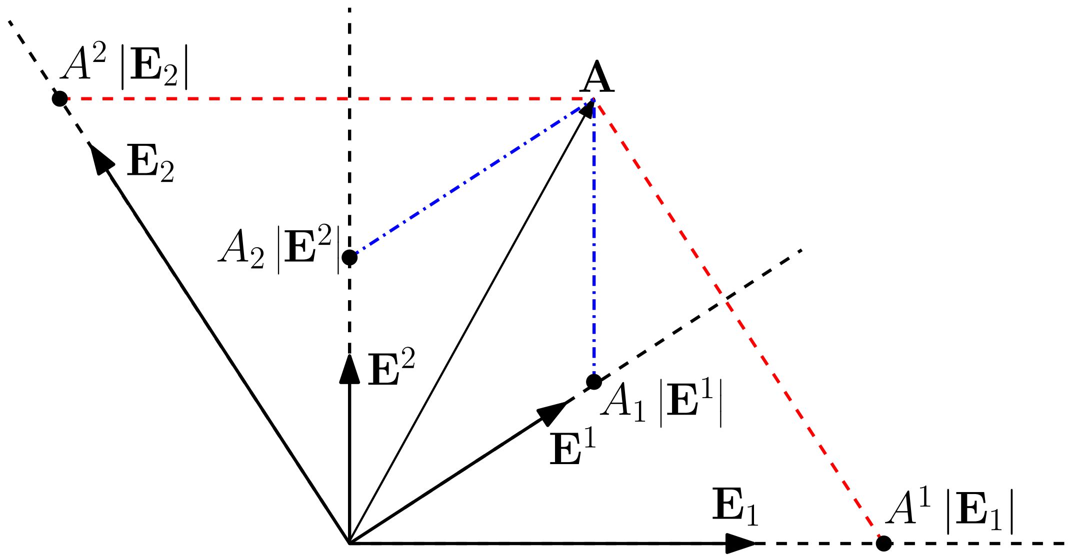


Figure 15: The representation of a vector **A** in covariant and contravariant basis vector sets in a 2D space where the components shown at the four points are with reference to unit vectors in the given directions, e.g. $A^1 |\mathbf{E}_1|$ is a component with reference to a unit vector in the direction of $\mathbf{E}_1$.

tensors are proper or ordinary tensors and hence they are **invariant** under coordinate transformations, while pseudo tensors are not proper tensors since they do not transform invariantly as they acquire a **minus** sign under improper orthogonal transformations which involve **inversion** of coordinate axes through the origin of coordinates with a change of system handedness.

Figure 16 demonstrates the behavior of a true vector **v** and a pseudo vector **p** where the former **keeps** its direction following a **reflection** of the coordinate system through the origin of coordinates while the latter **reverses** its direction following this operation.

Because true and pseudo tensors have different mathematical properties and represent different types of physical entities, the **terms** of consistent tensor expressions and equations should be **uniform** in their true and pseudo type, i.e. all terms are true or all are pseudo.

The **direct product** (refer to § 3.3) of true tensors is a **true** tensor. The direct product of **even** number of pseudo tensors is a **true** tensor, while the direct product of **odd**

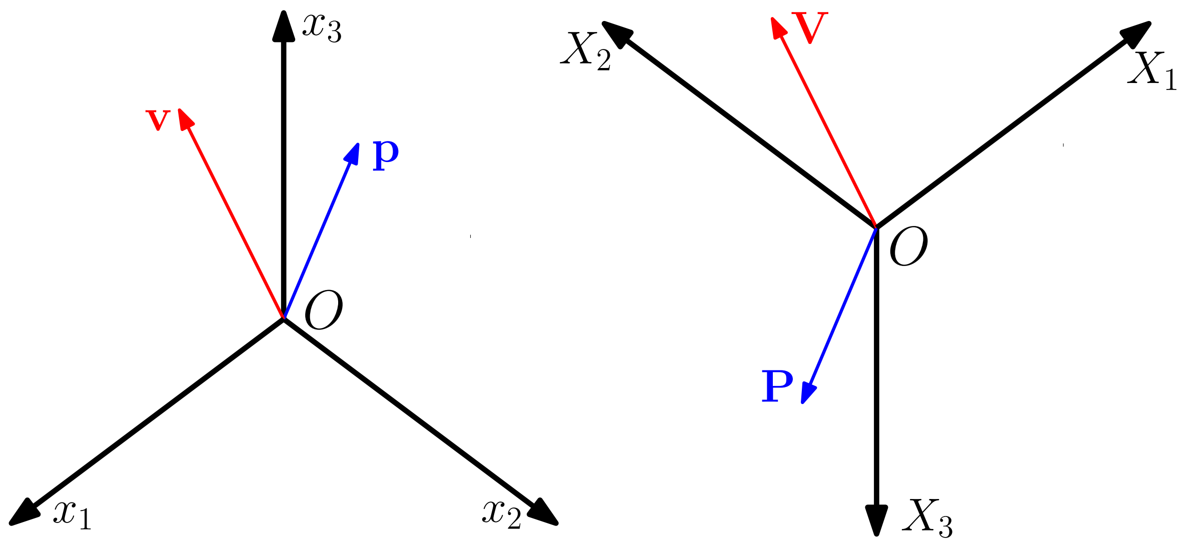


Figure 16: The behavior of a true vector (**v** and **V**) and a pseudo vector (**p** and **P**) on reflecting the coordinate system in the origin of coordinates. The lower case symbols stand for the objects in the original system while the upper case symbols stand for the same objects in the reflected system.

number of pseudo tensors is a **pseudo** tensor. The direct product of a mix of true and pseudo tensors is a true or pseudo tensor **depending** on the number of pseudo tensors involved in the product as being even or odd respectively.

Similar rules to those of the direct product apply to the **cross product**, including the **curl** operation, involving tensors (which are usually of rank-1) with the **addition** of a pseudo factor for each cross product operation. This factor is contributed by the **permutation** tensor ϵ which is implicit in the definition of the cross product (see Eqs. 173 and 192). As we will see in § 4.2, the permutation tensor is a pseudo tensor.

In summary, what determines the tensor type (true or pseudo) of the tensor terms involving direct[34] and cross products is the **parity** of the multiplicative factors of pseudo type **plus** the number of cross product operations involved since each cross product operation contributes an ϵ factor.

Examples of **true scalars** are temperature, mass and the dot product of two polar or

[34] Inner product (see § 3.5) is the result of a direct product (see § 3.3) operation followed by a contraction (see § 3.4) and hence it is like a direct product in this context.

two axial vectors, while examples of **pseudo scalars** are the dot product of an axial vector and a polar vector and the scalar triple product of polar vectors. Examples of **polar vectors** are displacement and acceleration, while examples of **axial vectors** are angular velocity and cross product of polar vectors in general, including the curl operation on polar vectors, due to the involvement of the permutation symbol ϵ which is a pseudo tensor as stated already. As indicated before, the essence of the distinction between true (i.e. polar) and pseudo (i.e. axial) vectors is that the direction of a pseudo vector depends on the observer choice of the handedness of the coordinate system whereas the direction of a true vector is independent of such a choice.

Examples of **true tensors of rank-2** are stress and rate of strain tensors, while examples of **pseudo tensors of rank-2** are direct products of two vectors: one polar and one axial. Examples of **true tensors** of **higher ranks** are piezoelectric moduli tensor (rank-3) and elasticity tensor (rank-4), while examples of **pseudo tensors** of **higher ranks** are the permutation tensor of these ranks.

2.6.3 Absolute and Relative Tensors

Considering an arbitrary transformation from a general coordinate system to another, a tensor of **weight** w is defined by the following general tensor transformation:

$$\bar{A}^{ij\ldots k}{}_{lm\ldots n} = \left|\frac{\partial x}{\partial \bar{x}}\right|^{w} \frac{\partial \bar{x}^i}{\partial x^a}\frac{\partial \bar{x}^j}{\partial x^b}\cdots\frac{\partial \bar{x}^k}{\partial x^c}\frac{\partial x^d}{\partial \bar{x}^l}\frac{\partial x^e}{\partial \bar{x}^m}\cdots\frac{\partial x^f}{\partial \bar{x}^n}A^{ab\ldots c}{}_{de\ldots f} \tag{83}$$

where $\left|\frac{\partial x}{\partial \bar{x}}\right|$ is the **Jacobian** of the transformation between the two systems.[35] When $w = 0$ the tensor is described as an **absolute** or true tensor, and when $w \neq 0$ the tensor is described as a **relative** tensor. When $w = -1$ the tensor may be described as a **pseudo** tensor, while when $w = 1$ the tensor may be described as a **tensor density**.[36] As indicated earlier, a tensor of m contravariant indices and n covariant indices may be described as a tensor of type (m, n). This may be extended to include the weight w as a third entry and hence the **type** of the tensor is identified by (m, n, w).

Relative tensors can be added and subtracted (see § 3.1) if they are of the same variance type and have the **same** weight;[37] the result is a tensor of the same type and weight. Also, relative tensors can be equated if they are of the same type and weight. Multiplication of relative tensors produces a relative tensor whose weight is the **sum** of the weights of the original tensors. Hence, if the weights are added up to a non-zero value the result is a relative tensor of that weight; otherwise it is an absolute tensor.

[35] The Jacobian J is the determinant of the Jacobian matrix $\mathbf{J}$ of the transformation between the unbarred and barred systems, that is:

$$J = \det(\mathbf{J}) = \begin{vmatrix} \frac{\partial x^1}{\partial \bar{x}^1} & \frac{\partial x^1}{\partial \bar{x}^2} & \cdots & \frac{\partial x^1}{\partial \bar{x}^n} \\ \frac{\partial x^2}{\partial \bar{x}^1} & \frac{\partial x^2}{\partial \bar{x}^2} & \cdots & \frac{\partial x^2}{\partial \bar{x}^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial \bar{x}^1} & \frac{\partial x^n}{\partial \bar{x}^2} & \cdots & \frac{\partial x^n}{\partial \bar{x}^n} \end{vmatrix} \tag{84}$$

For more details, the reader is advised to consult more advanced textbooks on this subject.

[36] Some of these labels are used differently by different authors as the terminology of tensor calculus is not universally approved and hence the conventions of each author should be checked. Also, there is an obvious overlap between this classification (i.e. absolute and relative) and the previous classification (i.e. true and pseudo) at least according to some conventions.

[37] This statement should be generalized by including $w = 0$ which corresponds to absolute tensors and hence "relative" in this statement is more general than being opposite to "absolute". Accordingly, and from the perspective of relative tensors (i.e. assuming that other qualifications such as matching in the indicial structure are met), two absolute tensors can be added/subtracted but an absolute and a relative tensor (i.e. with $w \neq 0$) cannot since they are "relative" tensors with different weights.

2.6.4 Isotropic and Anisotropic Tensors

Isotropic tensors are characterized by the property that the **values** of their components are **invariant** under coordinate transformation by **proper rotation** of axes. In contrast, the values of the components of anisotropic tensors are dependent on the orientation of the coordinate axes. Notable **examples** of isotropic tensors are scalars (rank-0), the vector **0** (rank-1), Kronecker delta δ_{ij} (rank-2) and Levi-Civita tensor ϵ_{ijk} (rank-3). Many tensors describing physical properties of materials, such as stress and magnetic susceptibility, are anisotropic.

Direct and **inner** products (see § 3.3 and 3.5) of isotropic tensors are isotropic tensors. The **zero** tensor of any rank is isotropic; therefore if the components of a tensor vanish in a particular coordinate system they will **vanish in all** properly and improperly rotated coordinate systems.[38] Consequently, if the components of two tensors are identical in a particular coordinate system they are **identical in all** transformed coordinate systems. This means that tensor equalities and identities are invariant under coordinate transformations, which is one of the main motivations for the use of tensors in mathematics and science. As indicated, all rank-0 tensors (scalars) are isotropic. Also, the zero vector, **0**, of any dimension is isotropic; in fact it is the only rank-1 isotropic tensor.

2.6.5 Symmetric and Anti-symmetric Tensors

These types of tensor apply to **high ranks** only (rank ≥ 2).[39] Moreover, these types are **not exhaustive**, even for tensors of rank ≥ 2, as there are high-rank tensors which are neither symmetric nor anti-symmetric. A **rank-2** tensor A_{ij} is **symmetric** *iff* for all i

[38] For improper rotation, this is more general than being isotropic.

[39] Symmetry and anti-symmetry of tensors require in their definition two free indices at least; hence a scalar with no index and a vector with a single index do not qualify to be symmetric or anti-symmetric.

and j the following condition is satisfied:

$$A_{ji} = A_{ij} \tag{85}$$

and **anti-symmetric** or **skew-symmetric** *iff* for all i and j the following condition is satisfied:

$$A_{ji} = -A_{ij} \tag{86}$$

Similar conditions apply to contravariant type tensors (refer also to the following).

A **rank-n** tensor $A_{i_1 \ldots i_n}$ is **symmetric** in its two indices i_j and i_l *iff* the following condition applies identically:

$$A_{i_1 \ldots i_l \ldots i_j \ldots i_n} = A_{i_1 \ldots i_j \ldots i_l \ldots i_n} \tag{87}$$

and **anti-symmetric** in its two indices i_j and i_l *iff* the following condition applies identically:

$$A_{i_1 \ldots i_l \ldots i_j \ldots i_n} = -A_{i_1 \ldots i_j \ldots i_l \ldots i_n} \tag{88}$$

Any **rank-2** tensor A_{ij} can be **synthesized** from (or decomposed into) a symmetric part $A_{(ij)}$, which is marked with **round brackets** enclosing the indices, and an anti-symmetric part $A_{[ij]}$, which is marked with **square brackets**, where the following relations apply:

$$A_{ij} = A_{(ij)} + A_{[ij]} \qquad A_{(ij)} = \frac{1}{2}\left(A_{ij} + A_{ji}\right) \qquad A_{[ij]} = \frac{1}{2}\left(A_{ij} - A_{ji}\right) \tag{89}$$

Similarly, a **rank-3** tensor A_{ijk} can be **symmetrized** by the following relation:

$$A_{(ijk)} = \frac{1}{3!}\left(A_{ijk} + A_{kij} + A_{jki} + A_{ikj} + A_{jik} + A_{kji}\right) \tag{90}$$

and **anti-symmetrized** by the following relation:

$$A_{[ijk]} = \frac{1}{3!}\left(A_{ijk} + A_{kij} + A_{jki} - A_{ikj} - A_{jik} - A_{kji}\right) \tag{91}$$

More generally, a **rank-n** tensor $A_{i_1 \ldots i_n}$ can be **symmetrized** by:

$$A_{(i_1 \ldots i_n)} = \frac{1}{n!}\,(\text{sum of all even \& odd permutations of indices } i\text{'s}) \tag{92}$$

and **anti-symmetrized** by:

$$A_{[i_1 \ldots i_n]} = \frac{1}{n!}\,(\text{sum of all even permutations minus sum of all odd permutations}) \tag{93}$$

A tensor of high rank (> 2) may be symmetrized or anti-symmetrized with respect to **only some** of its indices instead of all of its indices. For example, in the following the tensor $\mathbf{A}$ is symmetrized and anti-symmetrized only with respect to its first two indices:

$$A_{(ij)k} = \frac{1}{2}\left(A_{ijk} + A_{jik}\right) \qquad\qquad A_{[ij]k} = \frac{1}{2}\left(A_{ijk} - A_{jik}\right) \tag{94}$$

A tensor is described as **totally symmetric** *iff* it is symmetric with respect to all of its indices, that is:

$$A_{i_1 \ldots i_n} = A_{(i_1 \ldots i_n)} \tag{95}$$

and **totally anti-symmetric** *iff* it is anti-symmetric in all of its indices, that is:

$$A_{i_1 \ldots i_n} = A_{[i_1 \ldots i_n]} \tag{96}$$

For a totally anti-symmetric tensor, **non-zero entries** can occur only when all the indices are different.

It should be remarked that the indices whose exchange defines the symmetry and anti-symmetry relations should be of the **same variance type**, i.e. both upper or both lower. Another important remark is that the symmetry and anti-symmetry characteristic of a tensor is **invariant** under coordinate transformations. Hence, a symmetric/anti-symmetric tensor in one coordinate system is symmetric/anti-symmetric in all other coordinate systems. Similarly, a tensor which is neither symmetric nor anti-symmetric in one coordinate system remains so in all other coordinate systems.[40]

Finally, for a symmetric tensor A_{ij} and an anti-symmetric tensor B^{ij} (or the other way around) we have the following useful and widely used identity:

$$A_{ij}B^{ij} = 0 \tag{97}$$

This is because an exchange of indices will change the sign of one tensor only and this will change the sign of the term in the summation resulting in having a sum of terms which is identically zero due to the fact that each term in the sum has its own negation.

2.7 Exercises

2.1 Make a sketch of a rank-2 tensor A_{ij} in a 4D space similar to Figure 11. What this tensor looks like?

2.2 What are the two main types of notation used for labeling tensors? State two names for each.

2.3 Make a detailed comparison between the two types of notation in the previous question stating any advantages or disadvantages in using one of these notations or the other.

[40] In this context, like many other contexts in this book, there are certain restrictions on the type and conditions of the coordinate transformations under which such statements are valid. However, these details cannot be discussed here due to the elementary level of this book.

In which context each one of these notations is more appropriate to use than the other?

2.4 What is the principle of invariance of tensors and why it is one of the main reasons for the use of tensors in science?

2.5 What are the two different meanings of the term "covariant" in tensor calculus?

2.6 State the type of each one of the following tensors considering the number and position of indices (i.e. covariant, contravariant, rank, scalar, vector, etc.):

$$a^i \qquad B_i^{jk} \qquad f \qquad b_k \qquad C^{ji}$$

2.7 Define the following technical terms which are related to tensors: term, expression, equality, order, rank, zero tensor, unit tensor, free index, dummy index, covariant, contravariant, and mixed.

2.8 Which of the following is a scalar, vector or rank-2 tensor: temperature, stress, cross product of two vectors, dot product of two vectors, and rate of strain?

2.9 What is the number of entries of a rank-0 tensor in a 2D space and in a 5D space? What is the number of entries of a rank-1 tensor in these spaces?

2.10 What is the difference between the order and rank of a tensor considering the different conventions in this regard?

2.11 What is the number of entries of a rank-3 tensor in a 4D space? What is the number of entries of a rank-4 tensor in a 3D space?

2.12 Describe direct and inverse coordinate transformations between spaces and write the generic equations for these transformations.

2.13 What are proper and improper transformations? Draw a simple sketch to demonstrate them.

2.14 Define the following terms related to permutation of indices: permutation, even, odd, parity, and transposition.

2.15 Find all the permutations of the following four letters assuming no repetition: (i, j, k, l).

2.16 Give three even permutations and three odd permutations of the symbols $(\alpha, \beta, \gamma, \delta)$ in the stated order.

2.17 Discuss all the similarities and differences between free and dummy indices.

2.18 What is the maximum number of repetitive indices that can occur in each term of a legitimate tensor expression?

2.19 How many components are represented by each one of the following assuming a 4D space?

$$A_i^{jk} \qquad f + g \qquad C^{mn} - D^{nm} \qquad 5D_k + 4A_k = B_k$$

2.20 What is the "summation convention"? To what type of indices this convention applies?

2.21 Is it always the case that the summation convention applies when an index is repeated? If not, what precaution should be taken to avoid ambiguity and confusion?

2.22 In which cases a pair of dummy indices should be of different variance type (i.e. one upper and one lower)? In what type of coordinate systems these repeated indices can be of the same variance type and why?

2.23 What are the rules that the free indices should obey when they occur in the terms of tensor expressions and equalities?

2.24 What is illegitimate about the following tensor expressions and equalities considering in your answer all the possible violations?

$$A_{ij} + B^k_{ij} \qquad C^n - D_n = B_m \qquad A^i_j = A^j_i \qquad A_j = f$$

2.25 Which of the following tensor expressions and equalities is legitimate and which is illegitimate?

$$B^i + C^{ij}_j \qquad A_i - B^k_i \qquad C^m + D^m = B^m_{mm} \qquad B^i_k = A^i_k$$

State in each illegitimate case all the reasons for illegitimacy.

2.26 Which is right and which is wrong of the following tensor equalities?

$$\partial_n A_n = A_n \partial_n \qquad [\mathbf{B}]_k + [\mathbf{D}]_k = [\mathbf{B} + \mathbf{D}]_k \qquad ab = ba \qquad A^{ij} M_{kl} = M_{kl} A^{ji}$$

Explain in each case why the equality is right or wrong.

2.27 Choose from the Index six entries related to the general rules that apply to the mathematical expressions and equalities in tensor calculus.

2.28 Give at least two examples of tensors used in mathematics, science and engineering for each one of the following ranks: $0, 1, 2$ and 3.

2.29 State the special names given to the rank-0 and rank-1 tensors.

2.30 What is the difference, if any, between rank-2 tensors and matrices?

2.31 Is the following statement correct? If not, re-write it correctly: "all rank-0 tensors are vectors and vice versa, and all rank-1 tensors are scalars and vice versa".

2.32 Give clear and detailed definitions of scalars and vectors and compare them. What is common and what is different between the two?

2.33 Make a simple sketch of the nine unit dyads associated with the double directions of rank-2 tensors in a 3D space.

2.34 Name three of the scientific disciplines that heavily rely on tensor calculus notation and techniques.

2.35 What are the main features of tensor calculus that make it very useful and successful in mathematical, scientific and engineering applications.

2.36 Why tensor calculus is used in the formulation and presentation of the laws of physics?

2.37 Give concise definitions for the covariant and contravariant types of tensor.

2.38 Describe how the covariant and contravariant types are notated and how they differ in their transformation between coordinate systems.

2.39 Give examples of tensors used in mathematics and science which are covariant and other examples which are contravariant.

2.40 Write the mathematical transformation rules of the following tensors: A_{ijk} to $\bar{A}_{rst}$ and B^{mn} to $\bar{B}^{pq}$.

2.41 Explain how mixed type tensors are defined and notated in tensor calculus.

2.42 Write the mathematical rule for transforming the mixed type tensor $D^{ij}{}_{klm}$ to $\bar{D}^{pq}{}_{rst}$.

2.43 From the Index, find all the terms that start with the word "Mixed" and are related specifically to tensors of rank 2.

2.44 Express the following tensors in indicial notation: a rank-3 covariant tensor **A**, a rank-4 contravariant tensor **B**, a rank-5 mixed type tensor **C** which is covariant in ij indices and contravariant in kmn indices where the indices are ordered as $ikmnj$.

2.45 Write step-by-step, similar to the detailed example given in § 2.6.1, the mathematical transformations of the following tensors: A_{ij} to $\bar{A}_{rs}$, B^{lmn} to $\bar{B}^{pqr}$, $C^{ij}{}_{mn}$ to $\bar{C}^{pq}{}_{rs}$ and $D_m{}^{kl}$ to $\bar{D}_r{}^{st}$.

2.46 What is the relation between the rank and the (m, n) type of a tensor?

2.47 Write, in indicial notation, the following tensors: **A** of type $(0, 4)$, **B** of type $(3, 1)$, **C** of type $(0, 0)$, **D** of type $(3, 4)$, **E** of type $(2, 0)$ and **F** of type $(1, 1)$.

2.48 What is the rank of each one of the tensors in the previous question? Are there tensors among them which may not have been notated properly?

2.49 Which tensor provides the link between the covariant and contravariant types of a given tensor **D**?

2.50 What coordinate system(s) in which the covariant and contravariant types of a tensor do not differ? What is the usual tensor notation used in this case?

2.51 Define in detail, qualitatively and mathematically, the covariant and contravariant types of the basis vectors of a general coordinate system explaining all the symbols used in your definition.

2.52 Is it necessary that the basis vectors of the previous exercise are mutually orthogonal and/or of unit length?

2.53 Is the following statement correct? "A superscript in the denominator of partial derivatives is equivalent to a superscript in the numerator". Explain why.

2.54 What is the reciprocity relation that links the covariant and contravariant basis vectors? Express this relation mathematically.

2.55 What is the interpretation of the reciprocity relation (refer to Figure 14 in your explanation)?

2.56 Are the covariant and contravariant forms of a specific tensor **A** represent the same mathematical object? If so, in what sense they are equal from the perspective of different coordinate systems?

2.57 Correct, if necessary, the following statement: "A tensor of any rank (≥ 1) can be represented covariantly using contravariant basis tensors of that rank, or contravariantly using contravariant basis tensors, or in a mixed form using a mixed basis of the same type".

2.58 Make corrections, if needed, to the following equations assuming a general curvilinear coordinate system where, in each case, all the possible ways of correction should be considered:

$$\mathbf{B} = B^i \mathbf{E}^i \quad \mathbf{M} = M_{ij} \mathbf{E}^i \quad \mathbf{D} = D^i \mathbf{E}^i \mathbf{E}_j \quad \mathbf{C} = C^i \mathbf{E}_j \quad \mathbf{F} = F^n \mathbf{E}_n \quad \mathbf{T} = T^{rs} \mathbf{E}_s \mathbf{E}_r$$

2.59 What is the technical term used to label the following objects: $\mathbf{E}^i\mathbf{E}^j$, $\mathbf{E}_i\mathbf{E}_j$, $\mathbf{E}_i\mathbf{E}^j$ and $\mathbf{E}^i\mathbf{E}_j$? What they mean?

2.60 What sort of tensor components that the objects in the previous question should be associated with?

2.61 What is the difference between true and pseudo vectors? Which of these is called axial and which is called polar?

2.62 Make a sketch demonstrating the behavior of true and pseudo vectors.

2.63 Is the following statement correct? "The terms of tensor expressions and equations should be uniform in their true and pseudo type". Explain why.

2.64 There are four possibilities for the direct product of two tensors of true and pseudo types. Discuss all these possibilities with respect to the type of the tensor produced by this operation and if it is true or pseudo. Also discuss in detail the cross product and curl operations from this perspective.

2.65 Give examples for the true and pseudo types of scalars, vectors and rank-2 tensors.

2.66 Explain, in words and equations, the meaning of absolute and relative tensors. Do these intersect in some cases with true and pseudo tensors (at least according to some conventions)?

2.67 What "Jacobian" and "weight" mean in the context of absolute and relative tensors?

2.68 Someone stated: "$\mathbf{A}$ is a tensor of type $(2, 4, -1)$". What these three numbers refer to?

2.69 What is the type of the tensor in the previous exercise from the perspectives of lower and upper indices and absolute and relative tensors? What is the rank of this tensor?

2.70 What is the weight of a tensor $\mathbf{A}$ produced from multiplying a tensor of weight -1 by a tensor of weight 2? Is $\mathbf{A}$ relative or absolute? Is it true or not?

2.71 Define isotropic and anisotropic tensors and give examples for each using tensors of different ranks.

2.72 What is the state of the inner and outer products of two isotropic tensors?

2.73 Why if a tensor equation is valid in a particular coordinate system it should also be valid in all other coordinate systems under admissible coordinate transformations? Use the isotropy of the zero tensor in your explanation.

2.74 Define "symmetric" and "anti-symmetric" tensors and write the mathematical condition that applies to each assuming a rank-2 tensor.

2.75 Do we have symmetric/anti-symmetric scalars or vectors? If not, why?

2.76 Is it the case that any tensor of rank > 1 should be either symmetric or anti-symmetric?

2.77 Give an example, writing all the components in numbers or symbols, of a symmetric tensor of rank-2 in a 3D space. Do the same for an anti-symmetric tensor of the same rank.

2.78 Give, if possible, an example of a rank-2 tensor which is neither symmetric nor anti-symmetric assuming a 4D space.

2.79 Using the Index in the back of the book, gather all the terms related to the symmetric and anti-symmetric tensors including the symbols used in their notations.

2.80 Is it true that any rank-2 tensor can be decomposed into a symmetric part and an anti-symmetric part? If so, write down the mathematical expressions representing these parts in terms of the original tensor. Is this also true for a general rank-n tensor?

2.81 What is the meaning of the round and square brackets which are used to contain indices in the indexed symbol of a tensor (e.g. $A_{(ij)}$ and $B^{[km]n}$)?

2.82 Can the indices of symmetry/anti-symmetry be of different variance type?

2.83 Is it possible that a rank-n ($n > 2$) tensor is symmetric/anti-symmetric with respect to some, but not all, of its indices? If so, give an example of a rank-3 tensor which is symmetric or anti-symmetric with respect to only two of its indices.

2.84 For a rank-3 covariant tensor A_{ijk}, how many possibilities of symmetry and anti-symmetry do we have? Consider in your answer total, as well as partial, symmetry and anti-symmetry. Is there another possibility (i.e. the tensor is neither symmetric nor anti-symmetric with respect to any pair of its indices)?

2.85 Can a tensor be symmetric with respect to some combinations of its indices and anti-symmetric with respect to the other combinations? If so, can you give a simple example of such a tensor?

2.86 Repeat the previous exercise considering the additional possibility that the tensor is neither symmetric nor anti-symmetric with respect to another set of indices, i.e. it is symmetric, anti-symmetric and neither with respect to different sets of indices.[41]

2.87 **A** is a rank-3 totally symmetric tensor and **B** is a rank-3 totally anti-symmetric tensor. Write all the mathematical conditions that these tensors satisfy.

2.88 Justify the following statement: "For a totally anti-symmetric tensor, non-zero entries can occur only when all the indices are different". Use mathematical, as well as descriptive, language in your answer.

2.89 For a totally anti-symmetric tensor B_{ijk} in a 3D space, write all the elements of this tensor which are identically zero. Consider the possibility that it may be easier to find first the elements which are not identically zero, then exclude the rest.[42]

[41] The best way to tackle this sort of exercises is to build a table or array of appropriate dimensions where the indexed components in symbolic or numeric formats are considered and a trial and error approach is used to investigate the possibility of creating such a tensor.

[42] The concepts of repetitive and non-repetitive permutations may be useful in tackling this question.

Chapter 3

Tensor Operations

There are various operations that can be performed on tensors to produce other tensors in general. Examples of these operations are addition/subtraction, multiplication by a scalar (rank-0 tensor), multiplication of tensors (each of rank > 0), contraction and permutation. Some of these operations, such as addition and multiplication, involve **more than one tensor** while others, such as contraction and permutation, are performed on a **single tensor**. In this chapter we provide a glimpse on the main elementary tensor operations of **algebraic nature** that permeate tensor algebra and calculus.

First, we should remark that the last section of this chapter, which is about the quotient rule for tensor test, is added to this chapter because it is the most appropriate place for it in the present book considering the dependency of the definition of this rule on other tensor operations; otherwise the section is not about a tensor operation in the same sense as the operations presented in the other sections of this chapter. Another remark is that in tensor algebra division is allowed only for scalars, hence if the components of an indexed tensor should appear in a denominator, the tensor should be redefined to avoid this, e.g. $B_i = \frac{1}{A_i}$.

3.1 Addition and Subtraction

Tensors of the **same rank and type**[43] can be added algebraically to produce a tensor of the same rank and type, e.g.

$$a = b + c \qquad A_i = B_i - C_i \qquad A^i_j = B^i_j + C^i_j \tag{98}$$

The added/subtracted terms should have the same indicial structure with regard to their free indices, as explained in § 2.3; hence A^i_{jk} and B^j_{ik} cannot be added or subtracted although they are of the same rank and type, but A^{mi}_{mjk} and B^i_{jk} can be added and subtracted. Addition of tensors is **associative** and **commutative**, that is:[44]

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \tag{99}$$

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \tag{100}$$

3.2 Multiplication of Tensor by Scalar

A tensor can be multiplied by a scalar, which generally should **not be zero**, to produce a tensor of the same variance type, rank and indicial structure, e.g.

$$A^j_{ik} = aB^j_{ik} \tag{101}$$

where a is a non-zero scalar. As indicated by the equation, multiplying a tensor by a scalar means multiplying **each component** of the tensor by that scalar. Multiplication

[43] Here, "type" refers to variance type (covariant/contravariant/mixed) and true/pseudo type as well as other qualifications to which the tensors participating in an addition or subtraction operation should match such as having the same weight if they are relative tensors, as outlined previously (refer for example to § 2.6.3).

[44] Associativity and commutativity can include subtraction if the minus sign is absorbed in the subtracted tensor; in which case the operation is converted to addition.

by a scalar is **commutative**, and **associative** when more than two factors are involved.

3.3 Tensor Multiplication

This may also be called **outer** or **exterior** or **direct** or **dyadic multiplication**, although some of these names may be reserved for operations on vectors. On multiplying each component of a tensor of rank r by each component of a tensor of rank k, both of dimension m, a tensor of **rank** $(r+k)$ with m^{r+k} components is obtained where the **variance type** of each index (covariant or contravariant) is **preserved**.

For example, if **A** and **B** are covariant tensors of rank-1, then on multiplying **A** by **B** we obtain a covariant tensor **C** of rank-2 where the components of **C** are given by:

$$C_{ij} = A_i B_j \tag{102}$$

while on multiplying **B** by **A** we obtain a covariant tensor **D** of rank-2 where the components of **D** are given by:

$$D_{ij} = B_i A_j \tag{103}$$

Similarly, if **A** is a contravariant tensor of rank-2 and **B** is a covariant tensor of rank-2, then on multiplying **A** by **B** we obtain a mixed tensor **C** of rank-4 where the components of **C** are given by:

$$C^{ij}{}_{kl} = A^{ij} B_{kl} \tag{104}$$

while on multiplying **B** by **A** we obtain a mixed tensor **D** of rank-4 where the components of **D** are given by:

$$D_{ij}{}^{kl} = B_{ij} A^{kl} \tag{105}$$

In the outer product operation, it is generally understood that all the indices of the involved tensors have the same range although this may not always be the case.[45]

In general, the outer product of tensors **yields a tensor**. The outer product of a tensor of type (m, n) by a tensor of type (p, q) results in a tensor of **type** $(m + p, n + q)$. This means that the tensor rank in the outer product operation is additive and the operation conserves the variance type of each index of the tensors involved.

The direct multiplication of tensors may be marked by the symbol $\otimes$, mostly when using symbolic notation for tensors, e.g. $\mathbf{A} \otimes \mathbf{B}$. However, in the present book no symbol is being used for the operation of direct multiplication and hence the operation is symbolized by putting the symbols of the tensors side by side, e.g. $\mathbf{AB}$ where $\mathbf{A}$ and $\mathbf{B}$ are non-scalar tensors. In this regard, the reader should be vigilant to avoid confusion with the operation of matrix multiplication which, according to the notation of matrix algebra, is also symbolized as $\mathbf{AB}$ where $\mathbf{A}$ and $\mathbf{B}$ are matrices of compatible dimensions, since matrix multiplication is an inner product, rather than an outer product, operation.

The direct multiplication of tensors is **not commutative** in general as indicated above; however it is **distributive** with respect to the algebraic sum of tensors, that is:[46]

$$\mathbf{AB} \neq \mathbf{BA} \tag{106}$$

$$\mathbf{A}(\mathbf{B} \pm \mathbf{C}) = \mathbf{AB} \pm \mathbf{AC} \qquad\qquad (\mathbf{B} \pm \mathbf{C})\mathbf{A} = \mathbf{BA} \pm \mathbf{CA} \tag{107}$$

As indicated before, the rank-2 tensor constructed by the direct multiplication of two vectors is commonly called **dyad**. Tensors may be expressed as an outer product of vectors where the rank of the resultant product is equal to the number of the vectors involved, e.g. 2 for dyads and 3 for triads. However, not every tensor can be synthesized as a product of

[45] As indicated before, there are cases of tensors which are not uniformly dimensioned, and in some cases these tensors can be regarded as the result of an outer product of lower rank tensors.

[46] Regarding the associativity of direct multiplication, there seems to be cases in which this operation is not associative. The interested reader is advised to refer to the research literature on this subject.

lower rank tensors. Multiplication of a tensor by a scalar (refer to § 3.2) may be regarded as a special case of direct multiplication.

3.4 Contraction

The contraction operation of a tensor of rank > 1 is to make **two free indices identical**, by unifying their symbols, and perform summation over these repeated indices, e.g.

$$A_i^j \xrightarrow{\text{contraction}} A_i^i \tag{108}$$

$$A_{il}^{jk} \xrightarrow{\text{contraction on } jl} A_{im}^{mk} \tag{109}$$

Contraction results in a **reduction** of the rank by 2 since it implies the annihilation of two free indices. Therefore, the contraction of a rank-2 tensor is a scalar, the contraction of a rank-3 tensor is a vector, the contraction of a rank-4 tensor is a rank-2 tensor, and so on.

For general **non-Cartesian** coordinate systems, the pair of contracted indices should be **different** in their variance type, i.e. one upper and one lower. Hence, contraction of a mixed tensor of type (m, n) will, in general, produce a tensor of **type** $(m-1, n-1)$. A tensor of type (p, q) can, therefore, have $p \times q$ possible contractions, i.e. one contraction for each combination of lower and upper indices.

A common example of contraction is the **dot product** operation on vectors (see § Dot Product) which can be regarded as a direct multiplication (refer to § 3.3) of the two vectors, which results in a rank-2 tensor, followed by a contraction. Also, in matrix algebra, taking the **trace** of a square matrix, by summing its diagonal elements, can be considered as a contraction operation on the rank-2 tensor represented by the matrix, and hence it yields the trace which is a scalar.

Conducting a contraction operation on a tensor **results into a tensor**. Similarly, the application of a contraction operation on a relative tensor (see § 2.6.3) produces a relative tensor of the **same weight** as the original tensor.

3.5 Inner Product

On taking the **outer product** (refer to § 3.3) of two tensors of rank ≥ 1 followed by a **contraction** (refer to § 3.4) on two indices of the product, an inner product of the two tensors is formed. Hence, if one of the original tensors is of rank-m and the other is of rank-n, the inner product will be of **rank-**$(m+n-2)$. In the symbolic notation of tensor calculus, the inner product operation is usually symbolized by a **single dot** between the two tensors, e.g. $\mathbf{A} \cdot \mathbf{B}$, to indicate the contraction operation which follows the outer multiplication.

In general, the inner product is **not commutative**. When one[47] or both of the tensors involved in the inner product are of rank > 1 then the order of the multiplicands does matter in general, that is:[48]

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A} \tag{110}$$

However, the inner product operation is **distributive** with respect to the algebraic sum of tensors, that is:

$$\mathbf{A} \cdot (\mathbf{B} \pm \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} \pm \mathbf{A} \cdot \mathbf{C} \qquad\qquad (\mathbf{B} \pm \mathbf{C}) \cdot \mathbf{A} = \mathbf{B} \cdot \mathbf{A} \pm \mathbf{C} \cdot \mathbf{A} \tag{111}$$

[47] The non-commutativity of the inner product in the case of only one of the involved tensors is of rank > 1 may not be obvious; however, a simple example is multiplying A_i and $B_j^{\ kl}$ with a contraction of j and k or j and l.

[48] In fact, this statement is rather vague and rudimentary and may not apply in some cases. There are many details related to the issue of commutativity of inner product of tensors which cannot be discussed here due to the level of this book. In general, several issues should be considered in this regard such as the order of the indices in the outer product of the two tensors involved in the inner product, the (m, n) type of the outer product and whether the contracted indices are contributed by the same tensor or the two tensors involved in the product assuming that the first case is conventionally an inner product operation. Another important issue to be considered is that the contracted indices must, in general, be of opposite variance type. Many of these details can be worked out rather easily by the vigilant reader from first principles if they cannot be obtained from the textbooks of tensor calculus.

As indicated before (see § 3.4), the **dot product** of two vectors is an example of the inner product of tensors, i.e. it is an inner product of two rank-1 tensors to produce a rank-0 tensor. For example, if **a** is a covariant vector and **b** is a contravariant vector, then their dot product can be depicted as follow:

$$[\mathbf{ab}]_i^{\ j} = a_i b^j \quad \xrightarrow{\text{contraction}} \quad \mathbf{a} \cdot \mathbf{b} = a_i b^i \tag{112}$$

Another common example, from linear algebra, of inner product is the **multiplication** of a matrix representing a rank-2 tensor, by a vector, which is a rank-1 tensor, to produce a vector. For example, if **A** is a rank-2 covariant tensor and **b** is a contravariant vector, then their inner product can be depicted, according to tensor calculus, as follow:[49]

$$[\mathbf{Ab}]_{ij}^{\ \ k} = A_{ij} b^k \quad \xrightarrow{\text{contraction on } jk} \quad [\mathbf{A} \cdot \mathbf{b}]_i = A_{ij} b^j \tag{113}$$

This operation is equivalent to the above mentioned operation of multiplying a matrix by a vector as defined in linear algebra.

The **multiplication** of two $n \times n$ matrices, as defined in linear algebra, to produce another $n \times n$ matrix is another example of inner product (see Eq. 171). In this operation, each one of the matrices involved in the multiplication, as well as the product itself, can represent a rank-2 tensor.

For tensors whose outer product produces a tensor of rank > 2 and type (m, n) where $m, n > 0$, **various** contraction operations between different pairs of indices of opposite variance type can occur and hence more than one inner product, which are different in general, can be defined. Moreover, when the outer product produces a tensor of rank > 3

[49] It should be emphasized that we are using the symbolic notation of tensor calculus, rather than the matrix notation, in writing **Ab** and $\mathbf{A} \cdot \mathbf{b}$ to represent, respectively, the outer and inner products. In matrix notation, **Ab** is used to represent the product of a matrix by a vector which is an inner product according to the terminology of tensor calculus.

and type (m, n) where $m, n > 1$, **more than one** contraction can take place **simultaneously**.[50]

There are more **specialized types** of inner product; some of these may be defined differently by different authors. For example, a double inner product of two rank-2 tensors, $\mathbf{A}$ and $\mathbf{B}$, may be defined and denoted by double vertically- or horizontally-aligned dots (e.g. $\mathbf{A} : \mathbf{B}$ or $\mathbf{A} \cdot\cdot \mathbf{B}$) to indicate **double contraction** taking place between different pairs of indices. An instance of these types is the inner product with double contraction of **two dyads** which is commonly defined by:[51]

$$\mathbf{ab} : \mathbf{cd} = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) \tag{114}$$

The single dots in the right hand side of this equation symbolize the conventional dot product of two vectors. The result of this operation is obviously a scalar since it is the product of two scalars, as seen from the right hand side of the equation.

Some authors may define a different type of double contraction inner product of two dyads, symbolized by two horizontally-aligned dots, which may be called a **transposed contraction**, and is given by:

$$\mathbf{ab} \cdot\cdot \mathbf{cd} = \mathbf{ab} : \mathbf{dc} = (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \tag{115}$$

where the result is also a scalar. However, different authors may have different conventions and hence one should be on the lookout for such differences.

For two **rank-2** tensors, the aforementioned double contraction inner products are sim-

[50] In these statements, we assume that the contracted indices can be contributed by the same tensor in the product as well as by the two tensors (i.e. one index from each tensor); otherwise more details are required. We are also assuming a general coordinate system where the contracted indices should be opposite in their variance type.

[51] This is also defined differently by some authors.

ilarly defined as in the case of two dyads, that is:

$$\mathbf{A} : \mathbf{B} = A_{ij} B_{ij} \qquad\qquad \mathbf{A} \cdot\cdot \mathbf{B} = A_{ij} B_{ji} \tag{116}$$

Inner products with higher multiplicities of contraction are similarly defined, and hence they can be regarded as trivial extensions of the inner products with lower contraction multiplicities.

The inner product of tensors **produces a tensor** because the inner product is an outer product operation followed by a contraction operation and both of these operations on tensors produce tensors, as stated before.

3.6 Permutation

A tensor may be obtained by **exchanging** the indices of another tensor. For example, A^i_{kj} is a permutation of the tensor A^i_{jk}. A common example of the permutation operation of tensors is the transposition of a matrix (refer to § Special Matrices) representing a rank-2 tensor since the first and second indices, which represent the rows and columns of the matrix, are exchanged in this operation.

It is obvious that tensor permutation applies only to tensors of **rank** > 1 since no exchange of indices can occur on a scalar with no index or on a vector with a single index. The collection of tensors obtained by permuting the indices of a reference tensor may be called **isomers**.

3.7 Tensor Test: Quotient Rule

Sometimes a tensor-like object may be suspected for being a tensor; in such cases a test based on what is called the "quotient rule"[52] can be used to clarify the situation. Accord-

[52] This should not be confused with the quotient rule of differentiation.

ing to this rule, if the inner product of a suspected tensor with a known tensor is a tensor then the suspect is a tensor. In more formal terms, if it is not known if **A** is a tensor but it is known that **B** and **C** are tensors; moreover it is known that the following relation holds true in all rotated (i.e. properly-transformed) coordinate frames:

$$A_{pq\ldots k\ldots m}B_{ij\ldots k\ldots n} = C_{pq\ldots mij\ldots n} \tag{117}$$

then **A** is a tensor. Here, **A**, **B** and **C** are respectively of ranks m, n and $(m+n-2)$, where the rank of **C** is reduced by 2 due to the contraction on k which can be any index of **A** and **B** independently.[53]

Testing for being a tensor can also be done by applying the **first principles** through direct substitution in the transformation equations of tensors to see if the alleged tensor satisfies the transformation rules or not. However, using the quotient rule is generally more convenient and requires less work. It is noteworthy that the quotient rule may be considered by some authors as a replacement for the division operation which is not defined for tensors.

3.8 Exercises

3.1 Give preliminary definitions of the following tensor operations: addition, multiplication by a scalar, tensor multiplication, contraction, inner product and permutation. Which of these operations involve a single tensor?

3.2 Give typical examples of addition/subtraction for rank-n $(0 \le n \le 3)$ tensors.

3.3 Is it possible to add two tensors of different ranks or different variance types? Is addition of tensors associative or commutative?

[53] We assume, of course, that the rules of contraction of indices, such as being of opposite variance type in the case of non-Cartesian coordinates, are satisfied in this operation.

3.4 Discuss, in detail, the operation of multiplication of a tensor by a scalar and compare it to the operation of tensor multiplication. Can we regard multiplying two scalars as an example of multiplying a tensor by a scalar?

3.5 What is the meaning of the term "outer product" and what are the other terms used to label this operation?

3.6 **C** is a tensor of rank-3 and **D** is a tensor of rank-2, what is the rank of their outer product **CD**? What is the rank of **CD** if it is subjected subsequently to a double contraction operation?

3.7 **A** is a tensor of type (m, n) and **B** is a tensor of type (s, t), what is the type of their direct product **AB**?

3.8 Discuss the operations of dot and cross product of two vectors (see § Dot Product and Cross Product) from the perspective of the outer product operation of tensors.

3.9 Collect from the Index all the terms related to the tensor operations of addition and permutation and classify these terms according to each operation giving a short definition of each.

3.10 Are the following two statements correct (make corrections if necessary)? "The outer multiplication of tensors is commutative but not distributive over sum of tensors" and "The outer multiplication of two tensors may produce a scalar".

3.11 What is contraction of tensor? How many free indices are consumed in a single contraction operation?

3.12 Is it possible that the contracted indices are of the same variance type? If so, what is the condition that should be satisfied for this to happen?

3.13 **A** is a tensor of type (m, n) where $m, n > 1$, what is its type after two contraction operations assuming a general coordinate system?

3.14 Does the contraction operation change the weight of a relative tensor?

3.15 Explain how the operation of multiplication of two matrices, as defined in linear algebra, involves a contraction operation. What is the rank of each matrix and what is the rank of the product? Is this consistent with the rule of reduction of rank by contraction?

3.16 Explain, in detail, the operation of inner product of two tensors and how it is related to the operations of contraction and outer product of tensors.

3.17 What is the rank and type of a tensor resulting from an inner product operation of a tensor of type (m, n) with a tensor of type (s, t)? How many possibilities do we have for this inner product considering the different possibilities of the embedded contraction operation?

3.18 Give an example of a commutative inner product of two tensors and another example of a non-commutative inner product.

3.19 Is the inner product operation distributive over algebraic addition of tensors?

3.20 Give an example from matrix algebra of inner product of tensors explaining in detail how the two are related.

3.21 Discuss specialized types of inner product operations that involve more than one contraction operation focusing in particular on the operations $\mathbf{A} : \mathbf{B}$ and $\mathbf{A} \cdot\cdot \mathbf{B}$ where **A** and **B** are two tensors of rank > 1.

3.22 A double inner product operation is conducted on a tensor of type $(1, 1)$ with a tensor of type $(1, 2)$. How many possibilities do we have for this operation? What is the rank and type of the resulting tensor? Is it covariant, contravariant or mixed?

3.23 Gather from the Index all the terms that refer to notations used in the operations of inner and outer product of tensors.

3.24 Assess the following statement considering the two meanings of the word "tensor" related to the rank: "Inner product operation of two tensors does not necessarily produce a tensor". Can this statement be correct in a sense and wrong in another?

3.25 What is the operation of tensor permutation and how it is related to the operation of transposition of matrices?

3.26 Is it possible to permute scalars or vectors and why?

3.27 What is the meaning of the term "isomers"?

3.28 Describe in detail the quotient rule and how it is used as a test for tensors.

3.29 Why the quotient rule is used instead of the standard transformation equations of tensors?

Chapter 4

δ and ϵ Tensors

In this chapter, we conduct a preliminary investigation about the δ and ϵ tensors and their properties and functions as well as the relation between them. These tensors are of particular importance in tensor calculus due to their distinctive properties and unique transformation attributes. They are **numerical tensors** with fixed components in all coordinate systems. The first is called **Kronecker delta** or **unit tensor**, while the second is called **Levi-Civita**,[54] **permutation**, **anti-symmetric** and **alternating tensor**.

4.1 Kronecker δ

The Kronecker δ is a **rank-2** tensor in all dimensions. It is defined as:

$$\delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \qquad (i, j = 1, 2, \ldots n) \tag{118}$$

where n is the space dimension, and hence it can be considered as the **identity matrix**. For example, in a 3D space the Kronecker δ tensor is given by:

$$[\delta_{ij}] = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{119}$$

[54] This name is usually used for the rank-3 tensor. Also some authors distinguish between the permutation tensor and the Levi-Civita tensor even for rank-3. Moreover, some of the common labels and descriptions of ϵ are more specific to rank-3.

The components of the covariant, contravariant and mixed types of this tensor are the **same**, that is:

$$\delta_{ij} = \delta^{ij} = \delta^i_j = \delta^j_i \tag{120}$$

The Kronecker δ tensor is **symmetric**, that is:

$$\delta_{ij} = \delta_{ji} \qquad\qquad \delta^{ij} = \delta^{ji} \tag{121}$$

where $i, j = 1, 2, \ldots, n$. Moreover, it is **conserved**[55] under all proper and improper coordinate transformations. Since it is **conserved** under proper transformations, it is an **isotropic** tensor.[56]

4.2 Permutation ϵ

The permutation tensor ϵ has a rank **equal** to the number of dimensions, and hence a rank-n permutation tensor has n^n components. The **rank-2** permutation tensor ϵ_{ij} is defined by:

$$\epsilon_{12} = 1 \qquad\qquad \epsilon_{21} = -1 \qquad\qquad \epsilon_{11} = \epsilon_{22} = 0 \tag{122}$$

Similarly, the **rank-3** permutation tensor ϵ_{ijk} is defined by:

$$\epsilon_{ijk} = \begin{cases} 1 & (i, j, k \text{ is even permutation of 1,2,3}) \\ -1 & (i, j, k \text{ is odd permutation of 1,2,3}) \\ 0 & (\text{repeated index}) \end{cases} \tag{123}$$

Figure 17 is a graphical illustration of the rank-3 permutation tensor ϵ_{ijk} while Figure

[55] "Conserved" means that the tensor keeps the values of its components following a coordinate transformation.

[56] Here, being conserved under all transformations is stronger than being isotropic as the former applies even under improper coordinate transformations while isotropy is restricted to proper transformations.

18, which may be used as a mnemonic device, demonstrates the **cyclic** nature of the three even permutations of the indices of the rank-3 permutation tensor and the three odd permutations of these indices assuming no repetition in indices. The three permutations in each case are obtained by starting from a given number in the cycle and rotating in the given direction to obtain the other two numbers in the permutation.

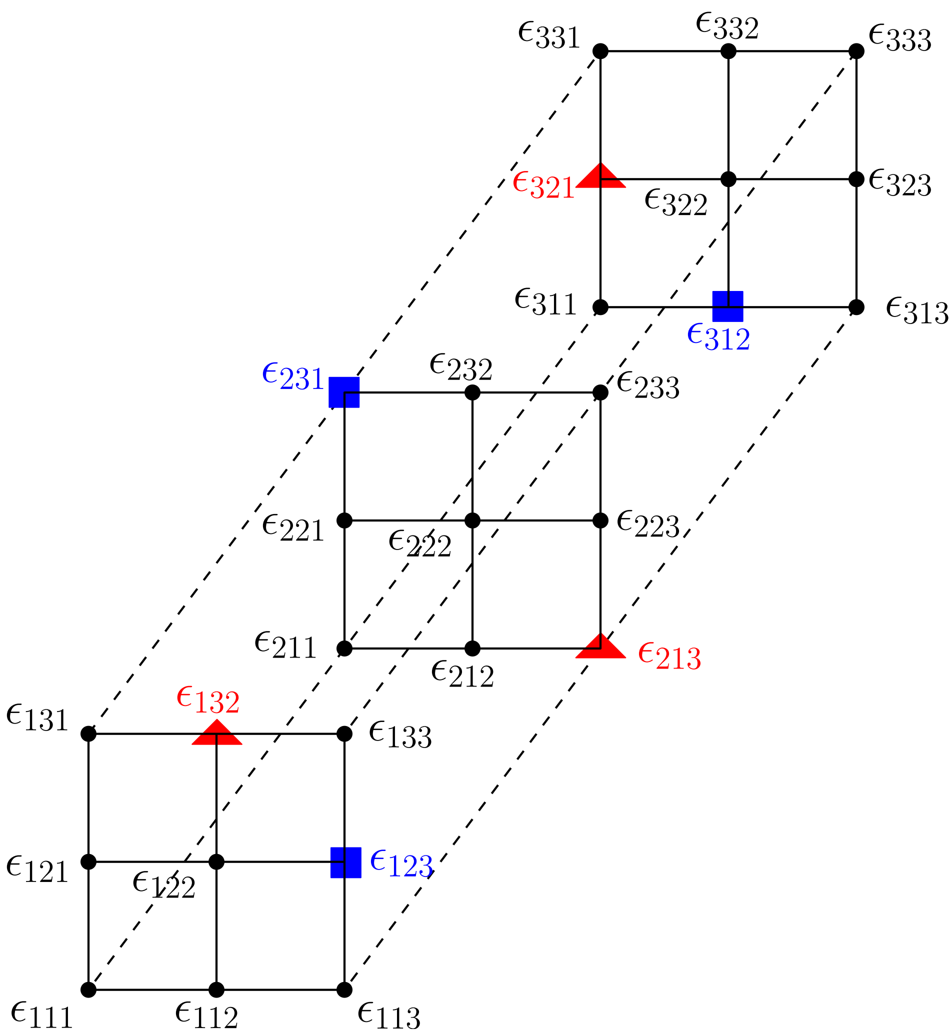


Figure 17: Graphical illustration of the rank-3 permutation tensor ϵ_{ijk} where circular nodes represent 0, square nodes represent 1 and triangular nodes represent -1.

The definition of the **rank-n** permutation tensor (i.e. $\epsilon_{i_1 i_2 \dots i_n}$) is similar to the definition of the rank-3 permutation tensor with regard to the repetition in its indices $(i_1, i_2, \cdots, i_n)$

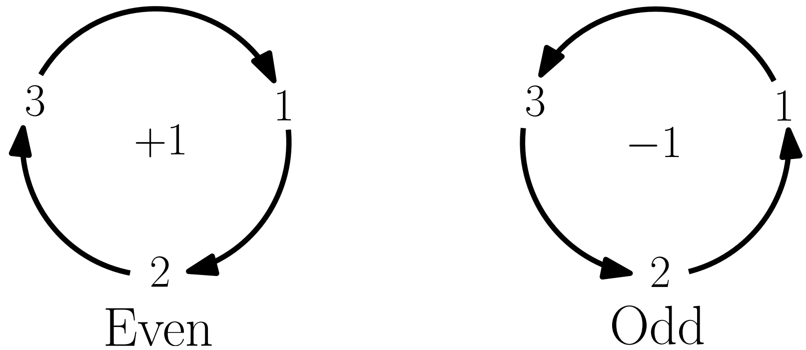


Figure 18: Graphical demonstration of the cyclic nature of the even and odd permutations of the indices of the rank-3 permutation tensor assuming no repetition in indices.

and being even or odd permutations in their correspondence to $(1, 2, \cdots, n)$, that is:

$$\epsilon_{i_1 i_2 \dots i_n} = \begin{cases} 1 & (i_1, i_2, \dots, i_n \text{ is even permutation of } 1, 2, \dots, n) \\ -1 & (i_1, i_2, \dots, i_n \text{ is odd permutation of } 1, 2, \dots, n) \\ 0 & \text{(repeated index)} \end{cases} \tag{124}$$

As well as the **inductive** definition of the permutation tensor (as given by Eqs. 122, 123 and 124), the permutation tensor of any rank can also be defined **analytically** where the entries of the tensor are calculated from closed form formulae. The entries of the **rank-2** permutation tensor can be calculated from the following **closed form** equation:

$$\epsilon_{ij} = (j - i) \tag{125}$$

Similarly, for the **rank-3** permutation tensor we have:

$$\epsilon_{ijk} = \frac{1}{2} (j - i) (k - i) (k - j) \tag{126}$$

while for the **rank-4** permutation tensor we have:

$$\epsilon_{ijkl} = \frac{1}{12}(j-i)(k-i)(l-i)(k-j)(l-j)(l-k) \tag{127}$$

More generally, the entries of the **rank-n** permutation tensor can be obtained from the following identity:

$$\epsilon_{a_1 a_2 \cdots a_n} = \prod_{i=1}^{n-1}\left[\frac{1}{i!}\prod_{j=i+1}^{n}(a_j - a_i)\right] = \frac{1}{S(n-1)}\prod_{1\le i<j\le n}(a_j - a_i) \tag{128}$$

where $S(n-1)$ is the **super factorial** function of the argument $(n-1)$ which is defined by:

$$S(k) = \prod_{i=1}^{k} i! = 1! \cdot 2! \cdot \ldots \cdot k! \tag{129}$$

A simpler formula for calculating the entries of the **rank-n** permutation tensor can be obtained from the previous one by dropping the magnitude of the multiplication factors and taking their **signs only**, that is:

$$\epsilon_{a_1 a_2 \cdots a_n} = \prod_{1\le i<j\le n}\operatorname{sgn}(a_j - a_i) = \operatorname{sgn}\left(\prod_{1\le i<j\le n}(a_j - a_i)\right) \tag{130}$$

where $\operatorname{sgn}(k)$ is the **sign function** of the argument k which is defined by:

$$\operatorname{sgn}(k) = \begin{cases} +1 & (k>0) \\ -1 & (k<0) \\ 0 & (k=0) \end{cases} \tag{131}$$

The permutation tensor is **totally anti-symmetric** (see § 2.6.5) in each pair of its indices, i.e. it changes sign on swapping any two of its indices, that is:

$$\epsilon_{i_1 \dots i_k \dots i_l \dots i_n} = -\epsilon_{i_1 \dots i_l \dots i_k \dots i_n} \tag{132}$$

The reason is that any exchange of two indices requires an even/odd number of single-step shifts to the right of the first index plus an odd/even number of single-step shifts to the left of the second index, so the total number of shifts is odd and hence it is an odd permutation of the original arrangement.

The permutation tensor is a **pseudo tensor** since it acquires a minus sign under an improper orthogonal transformation of coordinates, i.e. inversion of axes with possible superposition of rotation (see § 2.2). However, it is an **isotropic** tensor since it is conserved under proper coordinate transformations.

The permutation tensor may be considered as a **contravariant** relative tensor of **weight** $+1$ or a **covariant** relative tensor of **weight** -1. Hence, in 2D, 3D and nD spaces we have the following identities for the components of the permutation tensor:[57]

$$\epsilon_{ij} = \epsilon^{ij} \qquad \epsilon_{ijk} = \epsilon^{ijk} \qquad \epsilon_{i_1 i_2 \dots i_n} = \epsilon^{i_1 i_2 \dots i_n} \tag{133}$$

4.3 Useful Identities Involving δ or/and ϵ

In the following subsections we introduce and discuss a number of common identities which involve the Kronecker and permutations tensors. Some of these identities involve only one of these tensors while others involve both.

[57] We note that in equalities like this we are equating the components, as indicated above.

4.3.1 Identities Involving δ

When an index of the Kronecker delta is involved in a contraction operation by repeating an index in another tensor in its own term, the effect of this is to replace the shared index in the other tensor by the other index of the Kronecker delta, that is:

$$\delta_{ij} A_j = A_i \tag{134}$$

In such cases the Kronecker delta is described as an **index replacement** or **substitution operator**. Hence, we have:

$$\delta_{ij} \delta_{jk} = \delta_{ik} \tag{135}$$

Similarly:

$$\delta_{ij} \delta_{jk} \delta_{ki} = \delta_{ik} \delta_{ki} = \delta_{ii} = n \tag{136}$$

where n is the space dimension. The last part of this equation (i.e. $\delta_{ii} = n$) can be easily justified by the fact that δ_{ii} is the trace of the identity tensor considering the summation convention.

Due to the fact that the coordinates are **independent** of each other (see § 2.2), we also have the following identity:[58]

$$\frac{\partial x_i}{\partial x_j} = \partial_j x_i = x_{i,j} = \delta_{ij} \tag{137}$$

Hence, in an nD space we obtain the following identity from the last two identities:

$$\partial_i x_i = \delta_{ii} = n \tag{138}$$

[58] This identity, like many other identities in this chapter and in the book in general, is valid even for general coordinate systems although we use Cartesian notation to avoid unnecessary distraction at this level. The alert reader should be able to notate such identities in their general forms.

Based on the above identities and facts, the following identity can be shown to apply in orthonormal **Cartesian** coordinate systems:

$$\frac{\partial x_i}{\partial x_j} = \frac{\partial x_j}{\partial x_i} = \delta_{ij} = \delta^{ij} \tag{139}$$

This identity is based on the two facts that the coordinates are independent, and the covariant and contravariant types are the same in orthonormal Cartesian coordinate systems.

Similarly, for a coordinate system with a set of orthonormal[59] basis vectors, such as the orthonormal **Cartesian** system, the following identity can be easily proved:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \tag{140}$$

where the indexed **e** are the basis vectors. This identity is no more than a mathematical statement of the fact that the basis vectors in orthonormal systems are mutually orthogonal and of unit length.

Finally, the double inner product of two dyads (see § 3.5) formed by an orthonormal set of basis vectors of a given coordinate system satisfies the following identity:

$$\mathbf{e}_i\mathbf{e}_j : \mathbf{e}_k\mathbf{e}_l = \delta_{ik}\delta_{jl} \tag{141}$$

which is a combination of Eq. 114 and Eq. 140.

[59] As explained previously, "orthonormal" means that the vectors in the set are mutually orthogonal and each one of them is of unit length.

4.3.2 Identities Involving ϵ

From the definition of the rank-3 permutation tensor, we have the following identity which demonstrates the sense of **cyclic order** of the non-repetitive permutations of this tensor:

$$\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki} = -\epsilon_{ikj} = -\epsilon_{jik} = -\epsilon_{kji} \tag{142}$$

This identity is also a demonstration of the fact that the rank-3 permutation tensor is **totally anti-symmetric** in all of its indices since a shift of any two indices reverses its sign.[60] Moreover, it reflects the fact that this tensor has only **one independent non-zero component** since any one of the non-zero entries, all of which are given by Eq. 142, can be obtained from any other one of these entries.

We also have the following identity for the rank-n permutation tensor:[61]

$$\epsilon_{i_1 i_2 \cdots i_n} \epsilon_{i_1 i_2 \cdots i_n} = n! \tag{143}$$

This identity is based on the fact that the left hand side is actually the **sum of the squares** of $\epsilon_{i_1 i_2 \cdots i_n}$ over all the $n!$ non-repetitive permutations of n different indices where the value of ϵ of each one of these permutations is either $+1$ or -1 and hence in both cases their square is 1.

The double inner product of the **rank-3** permutation tensor and a **symmetric** tensor A_{jk} is given by the following identity:

$$\epsilon_{ijk} A_{jk} = 0 \tag{144}$$

[60] This may also be seen to apply to the zero entries of this tensor which correspond to the permutations with repetitive indices.

[61] This is a product of the rank-n permutation tensor by itself entry-by-entry with the application of the summation convention and hence it can be seen as a multi-contraction inner product of the permutation tensor by itself.

This is because an exchange of the two indices of A_{jk} does not affect its value due to the symmetry of A_{jk} whereas a similar exchange in these indices in ϵ_{ijk} results in a sign change; hence each term in the sum has its own negative and therefore the total sum is identically zero.

Another identity with a trivial outcome that involves the **rank-3** permutation tensor and a **vector A** is the following:

$$\epsilon_{ijk}A_iA_j = \epsilon_{ijk}A_iA_k = \epsilon_{ijk}A_jA_k = 0 \tag{145}$$

This can be explained by the fact that, due to the commutativity of ordinary multiplication, an exchange of the indices in A's will not affect the value but a similar exchange in the corresponding indices of ϵ_{ijk} will cause a change in sign; hence each term in the sum has its own negative and therefore the total sum will be zero.

Finally, for a set of **orthonormal basis vectors** in a 3D space with a right-handed coordinate system, the following identities are satisfied:

$$\mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ijk}\mathbf{e}_k \tag{146}$$

$$\mathbf{e}_i \cdot (\mathbf{e}_j \times \mathbf{e}_k) = \epsilon_{ijk} \tag{147}$$

These identities are based, respectively, on the forthcoming definitions of the cross product (see Eq. 173) and the scalar triple product (see Eq. 174) in tensor notation plus the fact that these vectors are unit vectors.

4.3.3 Identities Involving δ and ϵ

For the **rank-2** permutation tensor, we have the following identity which involves the Kronecker delta in 2D:

$$\epsilon_{ij}\epsilon_{kl} = \begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk} \tag{148}$$

This identity can simply be proved inductively by building a table for the values on the left and right hand sides as the indices are varied. The pattern of the indices in the determinant of this identity is simple, that is the indices of the first ϵ provide the indices for the rows while the indices of the second ϵ provide the indices for the columns.[62]

Another useful identity involving the **rank-2** permutation tensor with the Kronecker delta in 2D is the following:

$$\epsilon_{il}\epsilon_{kl} = \delta_{ik} \tag{149}$$

This can be obtained from the previous identity by replacing j with l followed by a minimal algebraic manipulation using tensor calculus rules.[63]

Similarly, we have the following identity which correlates the **rank-3** permutation tensor to the Kronecker delta in 3D:

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmn} &= \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix} = (\delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km}) - \\ & \qquad (\delta_{il}\delta_{jn}\delta_{km} + \delta_{im}\delta_{jl}\delta_{kn} + \delta_{in}\delta_{jm}\delta_{kl}) \end{aligned} \tag{150}$$

[62] The role of these indices in indexing the rows and columns can be shifted. This can be explained by the fact that the positions of the two epsilons can be exchanged, since ordinary multiplication is commutative, and hence the role of the epsilons in providing the indices for the rows and columns will be shifted. This can also be done by taking the transposition of the array of the determinant, which does not change the value of the determinant since $\det(\mathbf{A}) = \det(\mathbf{A}^T)$, with an exchange of the indices of the Kronecker symbols since the Kronecker symbol is symmetric in its two indices.

[63] That is:

$$\epsilon_{il}\epsilon_{kl} = \delta_{ik}\delta_{ll} - \delta_{il}\delta_{lk} = 2\delta_{ik} - \delta_{il}\delta_{lk} = 2\delta_{ik} - \delta_{ik} = \delta_{ik}$$

Again, the indices in the determinant of this identity follow the same pattern as that of Eq. 148.

Another useful identity in this category is the following:

$$\epsilon_{ijk}\epsilon_{lmk} = \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \tag{151}$$

This identity can be obtained from the identity of Eq. 150 by replacing n with k.[64] The pattern of the indices in this identity is as before if we exclude the repetitive indices.

More generally, the **determinantal form** of Eqs. 148 and 150, which link the rank-2 and rank-3 permutation tensors to the Kronecker tensors in 2D and 3D spaces, can be extended to link the **rank-n** permutation tensor to the Kronecker tensor in an nD space, that is:

$$\epsilon_{i_1 i_2 \cdots i_n}\,\epsilon_{j_1 j_2 \cdots j_n} = \begin{vmatrix} \delta_{i_1 j_1} & \delta_{i_1 j_2} & \cdots & \delta_{i_1 j_n} \\ \delta_{i_2 j_1} & \delta_{i_2 j_2} & \cdots & \delta_{i_2 j_n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{i_n j_1} & \delta_{i_n j_2} & \cdots & \delta_{i_n j_n} \end{vmatrix} \tag{152}$$

Again, the pattern of the indices in the determinant of this identity in their relation to the indices of the two epsilons follow the same rules as those of Eqs. 148 and 150.

The identity of Eq. 151, which may be called the **epsilon-delta identity**, the **contracted epsilon identity** or the **Levi-Civita identity**, is very useful in manipulating and simplifying tensor expressions and proving vector and tensor identities; examples of which will be seen in § 5.6. The sequence of indices of the δ's in the expanded form on the

[64] That is:

$$\begin{aligned} \epsilon_{ijk}\epsilon_{lmk} &= \delta_{il}\delta_{jm}\delta_{kk} + \delta_{im}\delta_{jk}\delta_{kl} + \delta_{ik}\delta_{jl}\delta_{km} - \delta_{il}\delta_{jk}\delta_{km} - \delta_{im}\delta_{jl}\delta_{kk} - \delta_{ik}\delta_{jm}\delta_{kl} \\ &= 3\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl} + \delta_{im}\delta_{jl} - \delta_{il}\delta_{jm} - 3\delta_{im}\delta_{jl} - \delta_{il}\delta_{jm} \\ &= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \end{aligned}$$

right hand side of this identity can be easily memorized using the following **mnemonic expression**:[65]

$$(FF \times SS) - (FS \times SF) \tag{153}$$

where the first and second F stand respectively for the first index in the first and second ϵ while the first and second S stand respectively for the second index in the first and second ϵ, as illustrated graphically in Figure 19. The mnemonic device of Eq. 153 can also be used to memorize the sequence of indices in Eq. 148.

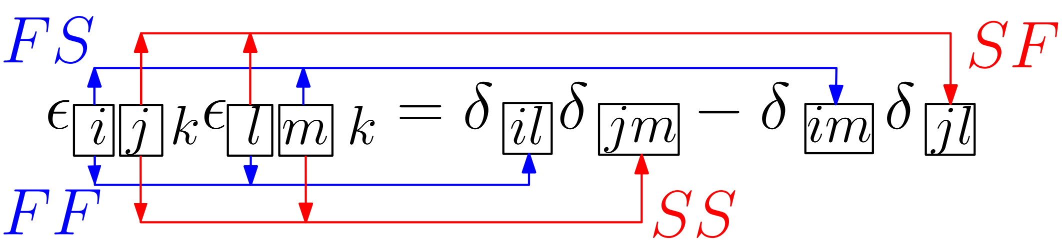


Figure 19: Graphical illustration of the mnemonic device of Eq. 153 which is used to remember the sequence of indices in the epsilon-delta identity of Eq. 151.

Other common identities in this category are:

$$\epsilon_{ijk}\epsilon_{ljk} = 2\delta_{il} \tag{154}$$

$$\epsilon_{ijk}\epsilon_{ijk} = 2\delta_{ii} = 2 \times 3 = 3! = 6 \tag{155}$$

The first of these identities can be obtained from Eq. 151 with the replacement of m with j followed by some basic tensor manipulation,[66] while the second can be obtained from

[65] In fact, the determinantal form given by Eq. 151 can also be considered as a mnemonic device where the first and second indices of the first ϵ index the rows while the first and second indices of the second ϵ index the columns, as given above. However, the mnemonic device of Eq. 153 is more economic in terms of the required work and more convenient in writing. It should also be remarked that the determinantal form in all the above equations is in fact a mnemonic device for these equations where the expanded form, if needed, can be easily obtained from the determinant which can be easily built following the simple pattern of indices, as explained above.

[66] That is:

$$\epsilon_{ijk}\epsilon_{ljk} = \delta_{il}\delta_{jj} - \delta_{ij}\delta_{jl} = 3\delta_{il} - \delta_{il} = 2\delta_{il}$$

the first by replacing l with i and applying the summation convention in 3D. The second identity is, in fact, an instance of Eq. 143 for a 3D space. Its restriction to the 3D space is justified by the fact that the rank and dimension of the permutation tensor are the **same**, which is 3 in this case. As indicated previously, δ_{ii} is the **trace** of the identity tensor, and hence in a 3D space it is equal to 3.

Another one of the common identities involving the **rank-3** permutation tensor with the Kronecker delta in 3D is the following:

$$\epsilon_{ijk}\delta_{1i}\delta_{2j}\delta_{3k} = \epsilon_{123} = 1 \tag{156}$$

This identity is based on the use of the Kronecker delta as an index replacement operator where each one of the deltas replaces an index in the permutation tensor.

Finally, the following identity can be obtained from the definition of the rank-3 permutation tensor (Eq. 123) and the use of the Kronecker delta as an index replacement operator (Eq. 134):

$$\epsilon_{ijk}\delta_{ij} = \epsilon_{ijk}\delta_{ik} = \epsilon_{ijk}\delta_{jk} = 0 \tag{157}$$

4.4 Generalized Kronecker δ

The generalized Kronecker delta is **defined inductively** by:

$$\delta^{i_1 \dots i_n}_{j_1 \dots j_n} = \begin{cases} 1 & (j_1 \dots j_n \text{ is even permutation of } i_1 \dots i_n) \\ -1 & (j_1 \dots j_n \text{ is odd permutation of } i_1 \dots i_n) \\ 0 & (\text{repeated } i\text{'s or } j\text{'s}) \end{cases} \tag{158}$$

It can also be **defined analytically** by the following $n \times n$ determinant:

$$\delta^{i_1 \dots i_n}_{j_1 \dots j_n} = \begin{vmatrix} \delta^{i_1}_{j_1} & \delta^{i_1}_{j_2} & \cdots & \delta^{i_1}_{j_n} \\ \delta^{i_2}_{j_1} & \delta^{i_2}_{j_2} & \cdots & \delta^{i_2}_{j_n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta^{i_n}_{j_1} & \delta^{i_n}_{j_2} & \cdots & \delta^{i_n}_{j_n} \end{vmatrix} \tag{159}$$

where the δ^i_j entries in the determinant are the **ordinary** Kronecker deltas as defined previously. In this equation, the **pattern** of the indices in the generalized Kronecker delta symbol $\delta^{i_1 \dots i_n}_{j_1 \dots j_n}$ in connection to the indices in the determinant is similar to the previous patterns, that is the upper indices in $\delta^{i_1 \dots i_n}_{j_1 \dots j_n}$ provide the upper indices in the ordinary deltas by indexing the rows of the determinant, while the lower indices in $\delta^{i_1 \dots i_n}_{j_1 \dots j_n}$ provide the lower indices in the ordinary deltas by indexing the columns of the determinant.

From the above given identities, it can be shown that:

$$\epsilon^{i_1 \dots i_n} \epsilon_{j_1 \dots j_n} = \begin{vmatrix} \delta^{i_1}_{j_1} & \delta^{i_1}_{j_2} & \cdots & \delta^{i_1}_{j_n} \\ \delta^{i_2}_{j_1} & \delta^{i_2}_{j_2} & \cdots & \delta^{i_2}_{j_n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta^{i_n}_{j_1} & \delta^{i_n}_{j_2} & \cdots & \delta^{i_n}_{j_n} \end{vmatrix} \tag{160}$$

Now, on comparing the last equation with the definition of the generalized Kronecker delta, i.e. Eq. 159, we conclude that:

$$\epsilon^{i_1 \dots i_n} \epsilon_{j_1 \dots j_n} = \delta^{i_1 \dots i_n}_{j_1 \dots j_n} \tag{161}$$

As an instance of Eq. 161, the relation between the rank-n permutation tensor in its covariant and contravariant forms and the generalized Kronecker delta in an nD space is given by:

$$\epsilon_{i_1 \dots i_n} = \delta^{1 \dots n}_{i_1 \dots i_n} \qquad\qquad \epsilon^{i_1 \dots i_n} = \delta^{i_1 \dots i_n}_{1 \dots n} \tag{162}$$

where the first of these equations can be obtained from Eq. 161 by substituting $(1 \dots n)$ for $(i_1 \dots i_n)$ in the two sides with relabeling j with i and noting that $\epsilon^{1 \dots n} = 1$, while the second equation can be obtained from Eq. 161 by substituting $(1 \dots n)$ for $(j_1 \dots j_n)$ and noting that $\epsilon_{1 \dots n} = 1$.

Hence, the permutation tensor ϵ can be considered as an **instance** of the generalized Kronecker delta. Consequently, the rank-n permutation tensor can be written as an $n \times n$ determinant consisting of the ordinary Kronecker deltas. Moreover, Eq. 162 can provide another definition for the permutation tensor in its covariant and contravariant forms, in addition to the previous inductive and analytic definitions of this tensor as given by Eqs. 124 and 128.

Returning to the widely used **epsilon-delta identity** of Eq. 151, if we define:[67]

$$\delta^{ij}_{lm} = \delta^{ijk}_{lmk} \tag{163}$$

and consider the above identities which correlate the permutation tensor, the generalized Kronecker tensor and the ordinary Kronecker tensor, then an identity equivalent to Eq. 151 that involves only the **generalized** and **ordinary** Kronecker deltas can be obtained, that is:

$$\delta^{ij}_{lm} = \delta^i_l \delta^j_m - \delta^i_m \delta^j_l \tag{164}$$

The mnemonic device of Eq. 153 can also be used with this form of the identity with minimal adjustments to the meaning of the symbols involved.

Other identities involving the permutation tensor and the ordinary Kronecker delta can also be formulated in terms of the generalized Kronecker delta.

[67] In fact this can be obtained from the determinantal form of δ^{ijk}_{lmn} by relabeling n with k.

4.5 Exercises

4.1 What "numerical tensor" means in connection with the Kronecker δ and the permutation ϵ tensors?

4.2 State all the names used to label the Kronecker and permutation tensors.

4.3 What is the meaning of "conserved under coordinate transformations" in relation to the Kronecker and permutation tensors?

4.4 State the mathematical definition of the Kronecker δ tensor.

4.5 What is the rank of the Kronecker δ tensor in an nD space?

4.6 Write down the matrix representing the Kronecker δ tensor in a 3D space.

4.7 Is there any difference between the components of the covariant, contravariant and mixed types of the Kronecker δ tensor?

4.8 Explain how the Kronecker δ acts as an index replacement operator giving an example in a mathematical form.

4.9 How many mathematical definitions of the rank-n permutation tensor we have? State one of these definitions explaining all the symbols involved.

4.10 What is the rank of the permutation tensor in an nD space?

4.11 Make a graphical illustration of the array representing the rank-2 and rank-3 permutation tensors.

4.12 Is there any difference between the components of the covariant and contravariant types of the permutation tensor?

4.13 How the covariant and contravariant types of the permutation tensor are related to the concept of relative tensor?

4.14 State the distinctive properties of the permutation tensor.

4.15 How many entries the rank-3 permutation tensor has? How many non-zero entries it has? How many independent entries it has?

4.16 Is the permutation tensor true or pseudo and why?

4.17 State, in words, the cyclic property of the even and odd non-repetitive permutations of the rank-3 permutation tensor with a simple sketch to illustrate this property.

4.18 From the Index, find all the terms which are common to the Kronecker and permutation tensors.

4.19 Correct the following equations:

$$\delta_{ij}A_j = A_j \qquad \delta_{ij}\delta_{jk} = \delta_{jk} \qquad \delta_{ij}\delta_{jk}\delta_{ki} = n! \qquad x_{i,j} = \delta_{ii}$$

4.20 In what type of coordinate system the following equation applies?

$$\partial_i x_j = \partial_j x_i$$

4.21 Complete the following equation assuming a 4D space:

$$\partial_i x_i = ?$$

4.22 Complete the following equations where the indexed $\mathbf{e}$ are orthonormal basis vectors of a particular coordinate system:

$$\mathbf{e}_i \cdot \mathbf{e}_j = ? \qquad\qquad \mathbf{e}_i\mathbf{e}_j : \mathbf{e}_k\mathbf{e}_l = ?$$

4.23 Write down the equations representing the cyclic order of the rank-3 permutation tensor. What is the conclusion from these equations with regard to the symmetry or anti-symmetry of this tensor and the number of its independent non-zero components?

4.24 Write the analytical expressions of the rank-3 and rank-4 permutation tensors.

4.25 Collect from the Index all the terms from matrix algebra which have connection to the Kronecker δ.

4.26 Correct, if necessary, the following equations:

$$\epsilon_{i_1 \cdots i_n} \epsilon_{i_1 \cdots i_n} = n \quad \epsilon_{ijk} C_j C_k = 0 \quad \epsilon_{ijk} D_{jk} = 0 \quad \mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ijk} \mathbf{e}_j \quad (\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_k = \epsilon_{ijk}$$

where $\mathbf{C}$ is a vector, $\mathbf{D}$ is a symmetric rank-2 tensor, and the indexed $\mathbf{e}$ are orthonormal basis vectors in a 3D space with a right-handed coordinate system.

4.27 What is wrong with the following equations?

$$\epsilon_{ijk} \delta_{1i} \delta_{2j} \delta_{3k} = -1 \qquad \epsilon_{ij} \epsilon_{kl} = \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \qquad \epsilon_{il} \epsilon_{kl} = \delta_{il} \qquad \epsilon_{ij} \epsilon_{ij} = 3!$$

4.28 Write the following in their determinantal form describing the general pattern of the relation between the indices of ϵ and δ and the indices of the rows and columns of the determinant:

$$\epsilon_{ijk} \epsilon_{lmk} \qquad \epsilon_{ijk} \epsilon_{lmn} \qquad \epsilon_{i_1 \cdots i_n} \epsilon_{j_1 \cdots j_n} \qquad \delta^{i_1 \ldots i_n}_{j_1 \ldots j_n}$$

4.29 Give two mnemonic devices used to memorize the widely used epsilon-delta identity and make a simple graphic illustration for one of these.

4.30 Correct, if necessary, the following equations:

$$\epsilon_{rst}\epsilon_{rst} = 3! \qquad \epsilon_{pst}\epsilon_{qst} = 2\delta_{pq} \qquad \epsilon_{rst}\delta_{rt} = \epsilon_{rst}\delta_{st}$$

4.31 State the mathematical definition of the generalized Kronecker delta $\delta^{i_1 \ldots i_n}_{j_1 \ldots j_n}$.

4.32 Write each one of $\epsilon_{i_1 \ldots i_n}$ and $\epsilon^{i_1 \ldots i_n}$ in terms of the generalized Kronecker δ.

4.33 Make a survey, based on the Index, about the general mathematical terms used in the operations conducted by using the Kronecker and permutation tensors.

4.34 Write the mathematical relation that links the covariant permutation tensor, the contravariant permutation tensor, and the generalized Kronecker delta.

4.35 State the widely used epsilon-delta identity in terms of the generalized and ordinary Kronecker deltas.

Chapter 5

Applications of Tensor Notation and Techniques

In this chapter, we provide common definitions in the language of tensor calculus for some basic concepts and operations from matrix and vector algebra. We also provide a preliminary investigation of the nabla based differential operators and operations using, in part, tensor notation. Common identities in vector calculus as well as the integral theorems of Gauss and Stokes are also presented from this perspective. Finally, we provide a rather extensive set of detailed examples about the use of tensor language and techniques in proving common mathematical identities from vector calculus.

5.1 Common Definitions in Tensor Notation

In this section, we give some common definitions of concepts and operations from vector and matrix algebra in tensor notation.

The **trace** of a matrix $\mathbf{A}$ representing a rank-2 tensor in an nD space is given by:

$$\mathrm{tr}\,(\mathbf{A}) = A_{ii} \qquad\qquad (i = 1, \ldots, n) \tag{165}$$

For a 3×3 matrix representing a rank-2 tensor in a 3D space, the **determinant** is given by:

$$\det(\mathbf{A}) = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} = \epsilon_{ijk} A_{1i} A_{2j} A_{3k} = \epsilon_{ijk} A_{i1} A_{j2} A_{k3} \tag{166}$$

where the last two equalities represent the expansion of the determinant by row and by column. Alternatively, the **determinant** of a 3×3 matrix can be given by:

$$\det(\mathbf{A}) = \frac{1}{3!} \epsilon_{ijk} \epsilon_{lmn} A_{il} A_{jm} A_{kn} \tag{167}$$

More generally, for an $n \times n$ matrix representing a rank-2 tensor in an nD space, the **determinant** is given by:

$$\begin{aligned} \det(\mathbf{A}) &= \epsilon_{i_1 \cdots i_n} A_{1 i_1} \ldots A_{n i_n} \\ &= \epsilon_{i_1 \cdots i_n} A_{i_1 1} \ldots A_{i_n n} \\ &= \frac{1}{n!} \epsilon_{i_1 \cdots i_n} \, \epsilon_{j_1 \cdots j_n} A_{i_1 j_1} \ldots A_{i_n j_n} \end{aligned} \tag{168}$$

The **inverse** of a 3×3 matrix $\mathbf{A}$ representing a rank-2 tensor is given by:

$$\left[\mathbf{A}^{-1}\right]_{ij} = \frac{1}{2 \det(\mathbf{A})} \epsilon_{jmn} \, \epsilon_{ipq} A_{mp} A_{nq} \tag{169}$$

The **multiplication** of a matrix $\mathbf{A}$ by a vector $\mathbf{b}$, as defined in linear algebra, is given by:

$$[\mathbf{Ab}]_i = A_{ij} b_j \tag{170}$$

It should be remarked that in the writing of $\mathbf{Ab}$ we are using **matrix notation**. According to the **symbolic notation** of tensors, the multiplication operation should be denoted by a **dot** between the symbols of the tensor and the vector, i.e. $\mathbf{A} \cdot \mathbf{b}$.[68]

[68] Matrix multiplication in matrix algebra is equivalent to inner product in tensor algebra.

The **multiplication** of two compatible matrices $\mathbf{A}$ and $\mathbf{B}$, as defined in linear algebra, is given by:

$$[\mathbf{AB}]_{ik} = A_{ij}B_{jk} \tag{171}$$

Again, we are using here matrix notation in the writing of $\mathbf{AB}$; otherwise a dot should be inserted between the symbols of the two matrices.

The **dot product** of two vectors of the same dimension is given by:

$$\mathbf{A} \cdot \mathbf{B} = \delta_{ij}A_iB_j = A_iB_i \tag{172}$$

The readers are referred to § 3.5 for a more general definition of this type of product that includes higher rank tensors. Similarly, the **cross product** of two vectors in a 3D space is given by:

$$[\mathbf{A} \times \mathbf{B}]_i = \epsilon_{ijk}A_jB_k \tag{173}$$

The **scalar triple product** of three vectors in a 3D space is given by:

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} = \epsilon_{ijk}A_iB_jC_k \tag{174}$$

while the **vector triple product** of three vectors in a 3D space is given by:

$$[\mathbf{A} \times (\mathbf{B} \times \mathbf{C})]_i = \epsilon_{ijk}\epsilon_{klm}A_jB_lC_m \tag{175}$$

The expression of the other principal form of the vector triple product [i.e. $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$] can be obtained from the above form by changing the order of the factors in the external cross product and reversing the sign; other operations, like relabeling the indices and exchanging some of the indices of the epsilons with a shift in sign, can then follow to obtain a more organized form. The expressions of the subsidiary forms of the vector triple product [e.g. $\mathbf{B}\times(\mathbf{A}\times\mathbf{C})$ or $(\mathbf{A}\times\mathbf{C})\times\mathbf{B}$] can be obtained from the above with relabeling the vectors in the indicial form according to their order in the symbolic form.

5.2 Scalar Invariants of Tensors

In the following, we list and write in tensor notation a number of **invariants** of low rank tensors which have special importance due to their widespread applications in vector and tensor calculus. All these invariants are **scalars**.

The value of a **scalar** (rank-0 tensor), which consists of a magnitude and a sign, is invariant under coordinate transformations. An invariant of a **vector** (rank-1 tensor) under coordinate transformations is its magnitude, i.e. length.[69] The main three independent scalar invariants of a **rank-2** tensor $\mathbf{A}$ under a change of basis are:

$$I = \mathrm{tr}(\mathbf{A}) = A_{ii} \tag{176}$$

$$II = \mathrm{tr}\left(\mathbf{A}^2\right) = A_{ij}A_{ji} \tag{177}$$

$$III = \mathrm{tr}\left(\mathbf{A}^3\right) = A_{ij}A_{jk}A_{ki} \tag{178}$$

Different forms of the three invariants of a rank-2 tensor $\mathbf{A}$, which are also widely used, are the following (noting that some of these definitions may belong to 3D specifically):

$$I_1 = I = A_{ii} \tag{179}$$

$$I_2 = \frac{1}{2}\left(I^2 - II\right) = \frac{1}{2}\left(A_{ii}A_{jj} - A_{ij}A_{ji}\right) \tag{180}$$

$$I_3 = \det(\mathbf{A}) = \frac{1}{3!}\left(I^3 - 3I\ II + 2III\right) = \frac{1}{3!}\epsilon_{ijk}\epsilon_{pqr}A_{ip}A_{jq}A_{kr} \tag{181}$$

where I, II and III are as defined in the previous set of equations.

[69] The direction is also invariant but it is not a scalar! In fact the magnitude alone is invariant under coordinate transformations even for pseudo vectors because it is a true scalar.

The invariants I, II and III can similarly be defined in terms of the invariants I_1, I_2 and I_3 as follow:

$$I = I_1 \tag{182}$$

$$II = I_1^2 - 2I_2 \tag{183}$$

$$III = I_1^3 - 3I_1I_2 + 3I_3 \tag{184}$$

Since the **determinant** of a matrix representing a rank-2 tensor is **invariant**, as given above, then if the determinant vanishes in one coordinate system, it will vanish in all transformed coordinate systems, and if not it will not. Consequently, if a rank-2 tensor is **invertible** in a particular coordinate system, it will be invertible in **all** coordinate systems, and if not it will not.

Ten joint invariants between two rank-2 tensors, $\mathbf{A}$ and $\mathbf{B}$, can be formed; these are: $\operatorname{tr}(\mathbf{A})$, $\operatorname{tr}(\mathbf{B})$, $\operatorname{tr}(\mathbf{A}^2)$, $\operatorname{tr}(\mathbf{B}^2)$, $\operatorname{tr}(\mathbf{A}^3)$, $\operatorname{tr}(\mathbf{B}^3)$, $\operatorname{tr}(\mathbf{A}\cdot\mathbf{B})$, $\operatorname{tr}(\mathbf{A}^2\cdot\mathbf{B})$, $\operatorname{tr}(\mathbf{A}\cdot\mathbf{B}^2)$ and $\operatorname{tr}(\mathbf{A}^2\cdot\mathbf{B}^2)$.

5.3 Common Differential Operations in Tensor Notation

In the first part of this section, we present the most common **nabla based** differential operations in **Cartesian** coordinates as defined by tensor notation. These operations are based on the various types of interaction between the vector differential operator nabla ∇ with tensors of different ranks where some of these interactions involve the dot and cross product operations. We then define in the subsequent parts of this section some of these differential operations in the most commonly used non-Cartesian coordinate systems, namely **cylindrical** and **spherical**, as well as in the **general orthogonal** coordinate systems.

Regarding the cylindrical and spherical systems, we could have used indexed generalized coordinates like x^1, x^2 and x^3 to represent the cylindrical coordinates (ρ, ϕ, z) and the spherical coordinates (r, θ, ϕ) and hence we express these operations in tensor notation as we did for the Cartesian system. However, for more clarity at this level and to follow the more conventional practice, we use the coordinates of these systems as suffixes in place of the usual indices used in the tensor notation.[70] Also, for clarity and simplicity of some formulae, we use numbers to index the coordinates of the general orthogonal system, where a 3D space is assumed, although this is not a compact way of representation since the formulae are given in their expanded form.

5.3.1 Cartesian Coordinate System

The **nabla** operator ∇ is a spatial partial differential operator which is defined in Cartesian coordinate systems by:

$$\nabla_i = \frac{\partial}{\partial x_i} \tag{185}$$

The **gradient** of a differentiable scalar function of position f is a vector obtained by applying the nabla operator on f and hence it is defined by:

$$[\nabla f]_i = \nabla_i f = \frac{\partial f}{\partial x_i} = \partial_i f = f_{,i} \tag{186}$$

Similarly, the **gradient** of a differentiable vector function of position $\mathbf{A}$ is the outer product (refer to § 3.3) between the ∇ operator and the vector and hence it is a rank-2 tensor given by:

$$[\nabla \mathbf{A}]_{ij} = \partial_i A_j \tag{187}$$

The **divergence** of a differentiable vector $\mathbf{A}$ is the **dot product** of the nabla operator

[70] There is another reason that is the components given in the cylindrical and spherical coordinates are physical, not covariant or contravariant, and hence suffixing with coordinates looks more appropriate. The interested reader should consult on this issue more advanced textbooks of tensor calculus.

and the vector **A** and hence it is a scalar given by:

$$\nabla \cdot \mathbf{A} = \delta_{ij} \frac{\partial A_i}{\partial x_j} = \frac{\partial A_i}{\partial x_i} = \nabla_i A_i = \partial_i A_i = A_{i,i} \tag{188}$$

The divergence operation can also be viewed as taking the gradient of the vector followed by a contraction. Hence, the divergence of a vector is invariant because it is the trace of a rank-2 tensor (see § 5.2).[71]

Similarly, the **divergence** of a differentiable rank-2 tensor **A** is a vector defined in one of its forms by:

$$[\nabla \cdot \mathbf{A}]_i = \partial_j A_{ji} \tag{189}$$

and in another form by:

$$[\nabla \cdot \mathbf{A}]_j = \partial_i A_{ji} \tag{190}$$

These two different forms can be given, respectively, in symbolic notation by:

$$\nabla \cdot \mathbf{A} \qquad\qquad \nabla \cdot \mathbf{A}^T \tag{191}$$

where $\mathbf{A}^T$ is the transpose of **A**.

More generally, the divergence of a tensor of rank $n \geq 2$, which is a tensor of rank-$(n-1)$, can be defined in several forms, which are different in general, depending on the combination of the contracted indices.

The **curl** of a differentiable vector **A** is the **cross product** of the nabla operator and the vector **A** and hence it is a vector defined by:

$$[\nabla \times \mathbf{A}]_i = \epsilon_{ijk} \frac{\partial A_k}{\partial x_j} = \epsilon_{ijk} \nabla_j A_k = \epsilon_{ijk} \partial_j A_k = \epsilon_{ijk} A_{k,j} \tag{192}$$

[71] It may also be argued more simply that the divergence of a vector is a scalar and hence it is invariant.

The curl operation may be generalized to tensors of rank > 1, and hence the **curl** of a differentiable rank-2 tensor $\mathbf{A}$ can be defined as a rank-2 tensor given by:

$$[\nabla \times \mathbf{A}]_{ij} = \epsilon_{imn} \partial_m A_{nj} \tag{193}$$

The **Laplacian** scalar operator acting on a differentiable scalar f is given by:

$$\Delta f = \nabla^2 f = \delta_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_i} = \nabla_{ii} f = \partial_{ii} f = f_{,ii} \tag{194}$$

The **Laplacian** operator acting on a differentiable vector $\mathbf{A}$ is defined for each component of the vector in a similar manner to the definition of the Laplacian acting on a scalar, that is:

$$\left[\nabla^2 \mathbf{A}\right]_i = \nabla^2 [\mathbf{A}]_i = \partial_{jj} A_i \tag{195}$$

The following **scalar differential operator** is commonly used in science (e.g. in fluid dynamics):

$$\mathbf{A} \cdot \nabla = A_i \nabla_i = A_i \frac{\partial}{\partial x_i} = A_i \partial_i \tag{196}$$

where $\mathbf{A}$ is a vector. As indicated earlier, the order of A_i and ∂_i should be respected. The following **vector differential operator** also has common applications in science:

$$[\mathbf{A} \times \nabla]_i = \epsilon_{ijk} A_j \partial_k \tag{197}$$

where, again, the order should be respected.

It should be remarked that the differentiation of a tensor **increases** its rank by one, by introducing an extra covariant index, unless it implies a contraction in which case it **reduces** the rank by one. Therefore the gradient of a scalar is a vector and the gradient of a vector is a rank-2 tensor $(\partial_i A_j)$, while the divergence of a vector is a scalar and the divergence of a rank-2 tensor is a vector ($\partial_j A_{ji}$ or $\partial_i A_{ji}$). This may be justified by the fact that nabla ∇ is a **vector operator**. On the other hand the Laplacian operator does not change the rank since it is a **scalar operator**; hence the Laplacian of a scalar is a scalar and the Laplacian of a vector is a vector.

5.3.2 Cylindrical Coordinate System

For the cylindrical system identified by the coordinates (ρ, ϕ, z) with a set of **orthonormal basis vectors** $\mathbf{e}_\rho$, $\mathbf{e}_\phi$ and $\mathbf{e}_z$ we have the following definitions for the nabla based operators and operations.[72]

The **nabla** operator ∇ is given by:

$$\nabla = \mathbf{e}_\rho \partial_\rho + \mathbf{e}_\phi \frac{1}{\rho} \partial_\phi + \mathbf{e}_z \partial_z \tag{198}$$

The **Laplacian** operator is given by:

$$\nabla^2 = \partial_{\rho\rho} + \frac{1}{\rho} \partial_\rho + \frac{1}{\rho^2} \partial_{\phi\phi} + \partial_{zz} \tag{199}$$

The **gradient** of a differentiable scalar f is given by:

$$\nabla f = \mathbf{e}_\rho \frac{\partial f}{\partial \rho} + \mathbf{e}_\phi \frac{1}{\rho} \frac{\partial f}{\partial \phi} + \mathbf{e}_z \frac{\partial f}{\partial z} \tag{200}$$

The **divergence** of a differentiable vector $\mathbf{A}$ is given by:

$$\nabla \cdot \mathbf{A} = \frac{1}{\rho} \left[\frac{\partial (\rho A_\rho)}{\partial \rho} + \frac{\partial A_\phi}{\partial \phi} + \frac{\partial (\rho A_z)}{\partial z} \right] \tag{201}$$

[72] It should be obvious that since ρ, ϕ and z are labels for specific coordinates and not variable indices, the summation convention does not apply.

The **curl** of a differentiable vector $\mathbf{A}$ is given by:

$$\nabla \times \mathbf{A} = \frac{1}{\rho} \begin{vmatrix} \mathbf{e}_\rho & \rho \mathbf{e}_\phi & \mathbf{e}_z \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \phi} & \frac{\partial}{\partial z} \\ A_\rho & \rho A_\phi & A_z \end{vmatrix} \tag{202}$$

For the **plane polar** coordinate system, these operators and operations can be obtained by **dropping** the z components or terms from the cylindrical form of the above operators and operations.

5.3.3 Spherical Coordinate System

For the spherical system identified by the coordinates (r, θ, ϕ) with a set of **orthonormal basis vectors** $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ we have the following definitions for the nabla based operators and operations.[73]

The **nabla** operator ∇ is given by:

$$\nabla = \mathbf{e}_r \partial_r + \mathbf{e}_\theta \frac{1}{r} \partial_\theta + \mathbf{e}_\phi \frac{1}{r \sin \theta} \partial_\phi \tag{203}$$

The **Laplacian** operator is given by:

$$\nabla^2 = \partial_{rr} + \frac{2}{r} \partial_r + \frac{1}{r^2} \partial_{\theta\theta} + \frac{\cos \theta}{r^2 \sin \theta} \partial_\theta + \frac{1}{r^2 \sin^2 \theta} \partial_{\phi\phi} \tag{204}$$

The **gradient** of a differentiable scalar f is given by:

$$\nabla f = \mathbf{e}_r \frac{\partial f}{\partial r} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial f}{\partial \theta} + \mathbf{e}_\phi \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \tag{205}$$

[73] Again, the summation convention does not apply to r, θ and ϕ.

The **divergence** of a differentiable vector $\mathbf{A}$ is given by:

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2 \sin\theta} \left[\sin\theta \frac{\partial \left(r^2 A_r\right)}{\partial r} + r \frac{\partial \left(\sin\theta A_\theta\right)}{\partial \theta} + r \frac{\partial A_\phi}{\partial \phi} \right] \tag{206}$$

The **curl** of a differentiable vector $\mathbf{A}$ is given by:

$$\nabla \times \mathbf{A} = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & r\sin\theta\mathbf{e}_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ A_r & rA_\theta & r\sin\theta A_\phi \end{vmatrix} \tag{207}$$

5.3.4 General Orthogonal Coordinate System

For general orthogonal systems in a 3D space identified by the coordinates (u^1, u^2, u^3) with a set of **unit basis vectors** $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ and **scale factors** h_1, h_2 and h_3 where $h_i = \left|\frac{\partial \mathbf{r}}{\partial u^i}\right|$ and $\mathbf{r} = x_i \mathbf{e}_i$ is the position vector, we have the following definitions for the nabla based operators and operations.

The **nabla** operator ∇ is given by:

$$\nabla = \frac{\mathbf{u}_1}{h_1} \frac{\partial}{\partial u^1} + \frac{\mathbf{u}_2}{h_2} \frac{\partial}{\partial u^2} + \frac{\mathbf{u}_3}{h_3} \frac{\partial}{\partial u^3} \tag{208}$$

The **Laplacian** operator is given by:

$$\nabla^2 = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u^1} \left(\frac{h_2 h_3}{h_1} \frac{\partial}{\partial u^1} \right) + \frac{\partial}{\partial u^2} \left(\frac{h_1 h_3}{h_2} \frac{\partial}{\partial u^2} \right) + \frac{\partial}{\partial u^3} \left(\frac{h_1 h_2}{h_3} \frac{\partial}{\partial u^3} \right) \right] \tag{209}$$

The **gradient** of a differentiable scalar f is given by:

$$\nabla f = \frac{\mathbf{u}_1}{h_1} \frac{\partial f}{\partial u^1} + \frac{\mathbf{u}_2}{h_2} \frac{\partial f}{\partial u^2} + \frac{\mathbf{u}_3}{h_3} \frac{\partial f}{\partial u^3} \tag{210}$$

The **divergence** of a differentiable vector **A** is given by:

$$\nabla \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u^1} (h_2 h_3 A_1) + \frac{\partial}{\partial u^2} (h_1 h_3 A_2) + \frac{\partial}{\partial u^3} (h_1 h_2 A_3) \right] \tag{211}$$

The **curl** of a differentiable vector **A** is given by:

$$\nabla \times \mathbf{A} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \mathbf{u}_1 & h_2 \mathbf{u}_2 & h_3 \mathbf{u}_3 \\ \frac{\partial}{\partial u^1} & \frac{\partial}{\partial u^2} & \frac{\partial}{\partial u^3} \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix} \tag{212}$$

5.4 Common Identities in Vector and Tensor Notation

In this section, we present some of the widely used identities of vector calculus using the traditional vector notation as well as its **equivalent** tensor notation. In the following bullet points, f and h are differentiable scalar fields; **A**, **B**, **C** and **D** are differentiable vector fields; and $\mathbf{r} = x_i \mathbf{e}_i$ is the position vector.

•

$$\begin{aligned} \nabla \cdot \mathbf{r} &= n \\ &\Updownarrow \\ \partial_i x_i &= n \end{aligned} \tag{213}$$

where n is the space dimension.

- $$\begin{aligned} \nabla \times \mathbf{r} &= \mathbf{0} \\ &\Updownarrow \\ \epsilon_{ijk}\partial_j x_k &= 0 \end{aligned} \tag{214}$$

- $$\begin{aligned} \nabla (\mathbf{a} \cdot \mathbf{r}) &= \mathbf{a} \\ &\Updownarrow \\ \partial_i (a_j x_j) &= a_i \end{aligned} \tag{215}$$

where **a** is a constant vector.

- $$\begin{aligned} \nabla \cdot (\nabla f) &= \nabla^2 f \\ &\Updownarrow \\ \partial_i (\partial_i f) &= \partial_{ii} f \end{aligned} \tag{216}$$

- $$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{A}) &= 0 \\ &\Updownarrow \\ \epsilon_{ijk}\partial_i \partial_j A_k &= 0 \end{aligned} \tag{217}$$

- $$\begin{aligned} \nabla \times (\nabla f) &= \mathbf{0} \\ &\Updownarrow \\ \epsilon_{ijk}\partial_j\partial_k f &= 0 \end{aligned} \tag{218}$$

- $$\begin{aligned} \nabla (fh) &= f\nabla h + h\nabla f \\ &\Updownarrow \\ \partial_i (fh) &= f\partial_i h + h\partial_i f \end{aligned} \tag{219}$$

- $$\begin{aligned} \nabla \cdot (f\mathbf{A}) &= f\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f \\ &\Updownarrow \\ \partial_i (fA_i) &= f\partial_i A_i + A_i\partial_i f \end{aligned} \tag{220}$$

- $$\begin{aligned} \nabla \times (f\mathbf{A}) &= f\nabla \times \mathbf{A} + \nabla f \times \mathbf{A} \\ &\Updownarrow \\ \epsilon_{ijk}\partial_j (fA_k) &= f\epsilon_{ijk}\partial_j A_k + \epsilon_{ijk} (\partial_j f) A_k \end{aligned} \tag{221}$$

- $$\begin{aligned} \mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\mathbf{C}\times\mathbf{A}) \\ \Updownarrow \qquad\qquad \Updownarrow \\ \epsilon_{ijk}A_iB_jC_k = \epsilon_{kij}C_kA_iB_j = \epsilon_{jki}B_jC_kA_i \end{aligned} \tag{222}$$

- $$\begin{aligned} \mathbf{A}\times(\mathbf{B}\times\mathbf{C}) &= \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B}) \\ &\Updownarrow \\ \epsilon_{ijk}A_j\epsilon_{klm}B_lC_m &= B_i(A_mC_m) - C_i(A_lB_l) \end{aligned} \tag{223}$$

- $$\begin{aligned} \mathbf{A}\times(\nabla\times\mathbf{B}) &= (\nabla\mathbf{B})\cdot\mathbf{A} - \mathbf{A}\cdot\nabla\mathbf{B} \\ &\Updownarrow \\ \epsilon_{ijk}\epsilon_{klm}A_j\partial_lB_m &= (\partial_iB_m)A_m - A_l(\partial_lB_i) \end{aligned} \tag{224}$$

- $$\begin{aligned} \nabla\times(\nabla\times\mathbf{A}) &= \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A} \\ &\Updownarrow \\ \epsilon_{ijk}\epsilon_{klm}\partial_j\partial_lA_m &= \partial_i(\partial_mA_m) - \partial_{ll}A_i \end{aligned} \tag{225}$$

- $$\begin{aligned} \nabla\left(\mathbf{A}\cdot\mathbf{B}\right) &= \mathbf{A}\times\left(\nabla\times\mathbf{B}\right)+\mathbf{B}\times\left(\nabla\times\mathbf{A}\right)+\left(\mathbf{A}\cdot\nabla\right)\mathbf{B}+\left(\mathbf{B}\cdot\nabla\right)\mathbf{A} \\ &\Updownarrow \\ \partial_i\left(A_mB_m\right) &= \epsilon_{ijk}A_j\left(\epsilon_{klm}\partial_lB_m\right)+\epsilon_{ijk}B_j\left(\epsilon_{klm}\partial_lA_m\right)+\left(A_l\partial_l\right)B_i+\left(B_l\partial_l\right)A_i \end{aligned} \tag{226}$$

- $$\begin{aligned} \nabla\cdot\left(\mathbf{A}\times\mathbf{B}\right) &= \mathbf{B}\cdot\left(\nabla\times\mathbf{A}\right)-\mathbf{A}\cdot\left(\nabla\times\mathbf{B}\right) \\ &\Updownarrow \\ \partial_i\left(\epsilon_{ijk}A_jB_k\right) &= B_k\left(\epsilon_{kij}\partial_iA_j\right)-A_j\left(\epsilon_{jik}\partial_iB_k\right) \end{aligned} \tag{227}$$

- $$\begin{aligned} \nabla\times\left(\mathbf{A}\times\mathbf{B}\right) &= \left(\mathbf{B}\cdot\nabla\right)\mathbf{A}+\left(\nabla\cdot\mathbf{B}\right)\mathbf{A}-\left(\nabla\cdot\mathbf{A}\right)\mathbf{B}-\left(\mathbf{A}\cdot\nabla\right)\mathbf{B} \\ &\Updownarrow \\ \epsilon_{ijk}\epsilon_{klm}\partial_j\left(A_lB_m\right) &= \left(B_m\partial_m\right)A_i+\left(\partial_mB_m\right)A_i-\left(\partial_jA_j\right)B_i-\left(A_j\partial_j\right)B_i \end{aligned} \tag{228}$$

- $$\begin{aligned} \left(\mathbf{A}\times\mathbf{B}\right)\cdot\left(\mathbf{C}\times\mathbf{D}\right) &= \begin{vmatrix} \mathbf{A}\cdot\mathbf{C} & \mathbf{A}\cdot\mathbf{D} \\ \mathbf{B}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{D} \end{vmatrix} \\ &\Updownarrow \\ \epsilon_{ijk}A_jB_k\epsilon_{ilm}C_lD_m &= \left(A_lC_l\right)\left(B_mD_m\right)-\left(A_mD_m\right)\left(B_lC_l\right) \end{aligned} \tag{229}$$

•

$$\begin{aligned}(\mathbf{A}\times\mathbf{B})\times(\mathbf{C}\times\mathbf{D}) &= [\mathbf{D}\cdot(\mathbf{A}\times\mathbf{B})]\,\mathbf{C}-[\mathbf{C}\cdot(\mathbf{A}\times\mathbf{B})]\,\mathbf{D}\\ &\Updownarrow\\ \epsilon_{ijk}\epsilon_{jmn}A_mB_n\epsilon_{kpq}C_pD_q &= (\epsilon_{qmn}D_qA_mB_n)\,C_i-(\epsilon_{pmn}C_pA_mB_n)\,D_i\end{aligned} \tag{230}$$

• In vector and tensor notations, the condition for a vector field **A** to be **solenoidal** is:

$$\begin{aligned}\nabla\cdot\mathbf{A} &= 0\\ &\Updownarrow\\ \partial_iA_i &= 0\end{aligned} \tag{231}$$

• In vector and tensor notations, the condition for a vector field **A** to be **irrotational** is:

$$\begin{aligned}\nabla\times\mathbf{A} &= \mathbf{0}\\ &\Updownarrow\\ \epsilon_{ijk}\partial_jA_k &= 0\end{aligned} \tag{232}$$

5.5 Integral Theorems in Tensor Notation

The **divergence theorem** for a differentiable **vector** field **A** in vector and tensor notations is given by:

$$\begin{aligned}\iiint_V \nabla\cdot\mathbf{A}\,d\tau &= \iint_S \mathbf{A}\cdot\mathbf{n}\,d\sigma\\ &\Updownarrow\\ \int_V \partial_iA_id\tau &= \int_S A_in_id\sigma\end{aligned} \tag{233}$$

where V is a bounded region in an nD space enclosed by a generalized surface S, $d\tau$ and $d\sigma$ are generalized volume and surface elements respectively, $\mathbf{n}$ and n_i are unit vector normal to the surface and its i^{th} component respectively, and the index i ranges over $1, \ldots, n$.

Similarly, the **divergence theorem** for a differentiable **rank-2** tensor field $\mathbf{A}$ in tensor notation for the first index is given by:

$$\int_V \partial_i A_{il} d\tau = \int_S A_{il} n_i d\sigma \tag{234}$$

while the **divergence theorem** for differentiable tensor fields of **higher rank A** in tensor notation for the index k is given by:

$$\int_V \partial_k A_{ij\ldots k\ldots m} d\tau = \int_S A_{ij\ldots k\ldots m} n_k d\sigma \tag{235}$$

Stokes theorem for a differentiable **vector** field $\mathbf{A}$ in vector and tensor notations is given by:

$$\begin{aligned} \iint_S (\nabla \times \mathbf{A}) \cdot \mathbf{n}\, d\sigma &= \int_C \mathbf{A} \cdot d\mathbf{r} \\ &\Updownarrow \\ \int_S \epsilon_{ijk} \partial_j A_k n_i d\sigma &= \int_C A_i dx_i \end{aligned} \tag{236}$$

where C stands for the perimeter of the surface S, and $d\mathbf{r}$ is a vector element tangent to the perimeter while the other symbols are as defined before.

Similarly, **Stokes theorem** for a differentiable **rank-2** tensor field $\mathbf{A}$ in tensor notation for the first index is given by:

$$\int_S \epsilon_{ijk} \partial_j A_{kl} n_i d\sigma = \int_C A_{il} dx_i \tag{237}$$

while **Stokes theorem** for differentiable tensor fields of **higher rank A** in tensor notation for the index k is given by:

$$\int_S \epsilon_{ijk}\partial_j A_{lm\ldots k\ldots n} n_i d\sigma = \int_C A_{lm\ldots k\ldots n} dx_k \tag{238}$$

5.6 Examples of Using Tensor Techniques to Prove Identities

In this section we provide some examples for using tensor techniques to prove vector and tensor identities. These examples, which are based on the identities given in § 5.4, demonstrate the elegance, efficiency and clarity of the methods and notation of tensor calculus.

- $\nabla \cdot \mathbf{r} = n$:

$$\begin{aligned} \nabla \cdot \mathbf{r} &= \partial_i x_i && \text{(Eq. 188)} \\ &= \delta_{ii} && \text{(Eq. 138)} \\ &= n && \text{(Eq. 138)} \end{aligned}$$

- $\nabla \times \mathbf{r} = \mathbf{0}$:

$$\begin{aligned} [\nabla \times \mathbf{r}]_i &= \epsilon_{ijk}\partial_j x_k && \text{(Eq. 192)} \\ &= \epsilon_{ijk}\delta_{kj} && \text{(Eq. 137)} \\ &= \epsilon_{ijj} && \text{(Eq. 134)} \\ &= 0 && \text{(Eq. 123)} \end{aligned}$$

Since i is a free index, the identity is proved for all components.

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

$$\begin{aligned} [\nabla (\mathbf{a} \cdot \mathbf{r})]_i &= \partial_i (a_j x_j) && \text{(Eqs. 186 \& 172)} \\ &= a_j \partial_i x_j + x_j \partial_i a_j && \text{(product rule)} \end{aligned}$$

$$
\begin{aligned}
&= a_j \partial_i x_j && (a_j \text{ is constant}) \\
&= a_j \delta_{ji} && (\text{Eq. } 137) \\
&= a_i && (\text{Eq. } 134) \\
&= [\mathbf{a}]_i && (\text{definition of index})
\end{aligned}
$$

Since i is a free index, the identity is proved for all components.

- $\nabla \cdot (\nabla f) = \nabla^2 f$:

$$
\begin{aligned}
\nabla \cdot (\nabla f) &= \partial_i [\nabla f]_i && (\text{Eq. } 188) \\
&= \partial_i (\partial_i f) && (\text{Eq. } 186) \\
&= \partial_i \partial_i f && (\text{rules of differentiation}) \\
&= \partial_{ii} f && (\text{definition of 2nd derivative}) \\
&= \nabla^2 f && (\text{Eq. } 194)
\end{aligned}
$$

- $\nabla \cdot (\nabla \times \mathbf{A}) = 0$:

$$
\begin{aligned}
\nabla \cdot (\nabla \times \mathbf{A}) &= \partial_i [\nabla \times \mathbf{A}]_i && (\text{Eq. } 188) \\
&= \partial_i (\epsilon_{ijk} \partial_j A_k) && (\text{Eq. } 192) \\
&= \epsilon_{ijk} \partial_i \partial_j A_k && (\partial \text{ not acting on } \epsilon) \\
&= \epsilon_{ijk} \partial_j \partial_i A_k && (\text{continuity condition}) \\
&= -\epsilon_{jik} \partial_j \partial_i A_k && (\text{Eq. } 142) \\
&= -\epsilon_{ijk} \partial_i \partial_j A_k && (\text{relabeling dummy indices } i \text{ and } j) \\
&= 0 && (\text{since } \epsilon_{ijk} \partial_i \partial_j A_k = -\epsilon_{ijk} \partial_i \partial_j A_k)
\end{aligned}
$$

This can also be concluded from line three by arguing that: since by the continuity condition ∂_i and ∂_j can change their order with no change in the value of the term while a corresponding change of the order of i and j in ϵ_{ijk} results in a sign change, we see that each term in the sum has its own negative and hence the terms add up to zero (see Eq. 145).

• $\nabla \times (\nabla f) = \mathbf{0}$:

$$
\begin{aligned}
[\nabla \times (\nabla f)]_i &= \epsilon_{ijk}\partial_j [\nabla f]_k && \text{(Eq. 192)} \\
&= \epsilon_{ijk}\partial_j (\partial_k f) && \text{(Eq. 186)} \\
&= \epsilon_{ijk}\partial_j\partial_k f && \text{(rules of differentiation)} \\
&= \epsilon_{ijk}\partial_k\partial_j f && \text{(continuity condition)} \\
&= -\epsilon_{ikj}\partial_k\partial_j f && \text{(Eq. 142)} \\
&= -\epsilon_{ijk}\partial_j\partial_k f && \text{(relabeling dummy indices } j \text{ and } k\text{)} \\
&= 0 && \text{(since } \epsilon_{ijk}\partial_j\partial_k f = -\epsilon_{ijk}\partial_j\partial_k f\text{)}
\end{aligned}
$$

This can also be concluded from line three by a similar argument to the one given in the previous point. Because $[\nabla \times (\nabla f)]_i$ is an arbitrary component, then each component is zero.

• $\nabla (fh) = f\nabla h + h\nabla f$:

$$
\begin{aligned}
[\nabla (fh)]_i &= \partial_i (fh) && \text{(Eq. 186)} \\
&= f\partial_i h + h\partial_i f && \text{(product rule)} \\
&= [f\nabla h]_i + [h\nabla f]_i && \text{(Eq. 186)} \\
&= [f\nabla h + h\nabla f]_i && \text{(Eq. 66)}
\end{aligned}
$$

Because i is a free index, the identity is proved for all components.

• $\nabla \cdot (f\mathbf{A}) = f\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f$:

$$
\nabla \cdot (f\mathbf{A}) = \partial_i [f\mathbf{A}]_i \qquad \text{(Eq. 188)}
$$

$$
\begin{aligned}
&= \partial_i (f A_i) && \text{(definition of index)} \\
&= f \partial_i A_i + A_i \partial_i f && \text{(product rule)} \\
&= f \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f && \text{(Eqs. 188 \& 196)}
\end{aligned}
$$

- $\nabla \times (f\mathbf{A}) = f\nabla \times \mathbf{A} + \nabla f \times \mathbf{A}$:

$$
\begin{aligned}
[\nabla \times (f\mathbf{A})]_i &= \epsilon_{ijk} \partial_j [f\mathbf{A}]_k && \text{(Eq. 192)} \\
&= \epsilon_{ijk} \partial_j (f A_k) && \text{(definition of index)} \\
&= f \epsilon_{ijk} \partial_j A_k + \epsilon_{ijk} (\partial_j f) A_k && \text{(product rule \& commutativity)} \\
&= f \epsilon_{ijk} \partial_j A_k + \epsilon_{ijk} [\nabla f]_j A_k && \text{(Eq. 186)} \\
&= [f \nabla \times \mathbf{A}]_i + [\nabla f \times \mathbf{A}]_i && \text{(Eqs. 192 \& 173)} \\
&= [f \nabla \times \mathbf{A} + \nabla f \times \mathbf{A}]_i && \text{(Eq. 66)}
\end{aligned}
$$

Because i is a free index, the identity is proved for all components.

- $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})$:

$$
\begin{aligned}
\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) &= \epsilon_{ijk} A_i B_j C_k && \text{(Eq. 174)} \\
&= \epsilon_{kij} A_i B_j C_k && \text{(Eq. 142)} \\
&= \epsilon_{kij} C_k A_i B_j && \text{(commutativity)} \\
&= \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) && \text{(Eq. 174)} \\
&= \epsilon_{jki} A_i B_j C_k && \text{(Eq. 142)} \\
&= \epsilon_{jki} B_j C_k A_i && \text{(commutativity)} \\
&= \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) && \text{(Eq. 174)}
\end{aligned}
$$

The negative permutations of this identity can be similarly obtained and proved by changing the order of the vectors in the cross products which results in a sign change.

• $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$:

$$
\begin{aligned}
[\mathbf{A} \times (\mathbf{B} \times \mathbf{C})]_i &= \epsilon_{ijk} A_j [\mathbf{B} \times \mathbf{C}]_k && \text{(Eq. 173)} \\
&= \epsilon_{ijk} A_j \epsilon_{klm} B_l C_m && \text{(Eq. 173)} \\
&= \epsilon_{ijk} \epsilon_{klm} A_j B_l C_m && \text{(commutativity)} \\
&= \epsilon_{ijk} \epsilon_{lmk} A_j B_l C_m && \text{(Eq. 142)} \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) A_j B_l C_m && \text{(Eq. 151)} \\
&= \delta_{il}\delta_{jm} A_j B_l C_m - \delta_{im}\delta_{jl} A_j B_l C_m && \text{(distributivity)} \\
&= (\delta_{il} B_l)(\delta_{jm} A_j C_m) - (\delta_{im} C_m)(\delta_{jl} A_j B_l) && \text{(commutativity \& grouping)} \\
&= B_i (A_m C_m) - C_i (A_l B_l) && \text{(Eq. 134)} \\
&= B_i (\mathbf{A} \cdot \mathbf{C}) - C_i (\mathbf{A} \cdot \mathbf{B}) && \text{(Eq. 172)} \\
&= [\mathbf{B}(\mathbf{A} \cdot \mathbf{C})]_i - [\mathbf{C}(\mathbf{A} \cdot \mathbf{B})]_i && \text{(definition of index)} \\
&= [\mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})]_i && \text{(Eq. 66)}
\end{aligned}
$$

Because i is a free index, the identity is proved for all components.

Other variants of this identity [e.g. $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$] can be obtained and proved similarly by changing the order of the factors in the external cross product with adding a minus sign.

• $\mathbf{A} \times (\nabla \times \mathbf{B}) = (\nabla \mathbf{B}) \cdot \mathbf{A} - \mathbf{A} \cdot \nabla \mathbf{B}$:

$$
\begin{aligned}
[\mathbf{A} \times (\nabla \times \mathbf{B})]_i &= \epsilon_{ijk} A_j [\nabla \times \mathbf{B}]_k && \text{(Eq. 173)} \\
&= \epsilon_{ijk} A_j \epsilon_{klm} \partial_l B_m && \text{(Eq. 192)} \\
&= \epsilon_{ijk} \epsilon_{klm} A_j \partial_l B_m && \text{(commutativity)} \\
&= \epsilon_{ijk} \epsilon_{lmk} A_j \partial_l B_m && \text{(Eq. 142)} \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) A_j \partial_l B_m && \text{(Eq. 151)} \\
&= \delta_{il}\delta_{jm} A_j \partial_l B_m - \delta_{im}\delta_{jl} A_j \partial_l B_m && \text{(distributivity)}
\end{aligned}
$$

$$
\begin{aligned}
&= A_m \partial_i B_m - A_l \partial_l B_i && \text{(Eq. 134)} \\
&= (\partial_i B_m) A_m - A_l (\partial_l B_i) && \text{(commutativity \& grouping)} \\
&= [(\nabla \mathbf{B}) \cdot \mathbf{A}]_i - [\mathbf{A} \cdot \nabla \mathbf{B}]_i && \text{(Eq. 187 \& § 3.5)} \\
&= [(\nabla \mathbf{B}) \cdot \mathbf{A} - \mathbf{A} \cdot \nabla \mathbf{B}]_i && \text{(Eq. 66)}
\end{aligned}
$$

Because i is a free index, the identity is proved for all components.

• $\nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$:

$$
\begin{aligned}
[\nabla \times (\nabla \times \mathbf{A})]_i &= \epsilon_{ijk} \partial_j [\nabla \times \mathbf{A}]_k && \text{(Eq. 192)} \\
&= \epsilon_{ijk} \partial_j (\epsilon_{klm} \partial_l A_m) && \text{(Eq. 192)} \\
&= \epsilon_{ijk} \epsilon_{klm} \partial_j (\partial_l A_m) && (\partial \text{ not acting on } \epsilon) \\
&= \epsilon_{ijk} \epsilon_{lmk} \partial_j \partial_l A_m && \text{(Eq. 142 \& definition of derivative)} \\
&= (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \partial_j \partial_l A_m && \text{(Eq. 151)} \\
&= \delta_{il} \delta_{jm} \partial_j \partial_l A_m - \delta_{im} \delta_{jl} \partial_j \partial_l A_m && \text{(distributivity)} \\
&= \partial_m \partial_i A_m - \partial_l \partial_l A_i && \text{(Eq. 134)} \\
&= \partial_i (\partial_m A_m) - \partial_{ll} A_i && (\partial \text{ shift, grouping \& Eq. 4)} \\
&= [\nabla (\nabla \cdot \mathbf{A})]_i - \left[\nabla^2 \mathbf{A}\right]_i && \text{(Eqs. 188, 186 \& 195)} \\
&= \left[\nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}\right]_i && \text{(Eq. 66)}
\end{aligned}
$$

Because i is a free index, the identity is proved for all components.

This identity can also be considered as an instance of the identity before the last one, observing that in the second term on the right hand side the Laplacian should precede the vector, and hence no independent proof is required.

• $\nabla (\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{A}$:

We start from the right hand side and end with the left hand side:

$$[\mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{A}]_i =$$

$$[\mathbf{A} \times (\nabla \times \mathbf{B})]_i + [\mathbf{B} \times (\nabla \times \mathbf{A})]_i + [(\mathbf{A} \cdot \nabla) \mathbf{B}]_i + [(\mathbf{B} \cdot \nabla) \mathbf{A}]_i =$$

(Eq. 66)

$$\epsilon_{ijk} A_j [\nabla \times \mathbf{B}]_k + \epsilon_{ijk} B_j [\nabla \times \mathbf{A}]_k + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(Eqs. 173, 188 & indexing)

$$\epsilon_{ijk} A_j (\epsilon_{klm} \partial_l B_m) + \epsilon_{ijk} B_j (\epsilon_{klm} \partial_l A_m) + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(Eq. 192)

$$\epsilon_{ijk} \epsilon_{klm} A_j \partial_l B_m + \epsilon_{ijk} \epsilon_{klm} B_j \partial_l A_m + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(commutativity)

$$\epsilon_{ijk} \epsilon_{lmk} A_j \partial_l B_m + \epsilon_{ijk} \epsilon_{lmk} B_j \partial_l A_m + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(Eq. 142)

$$(\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) A_j \partial_l B_m + (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) B_j \partial_l A_m + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(Eq. 151)

$$(\delta_{il} \delta_{jm} A_j \partial_l B_m - \delta_{im} \delta_{jl} A_j \partial_l B_m) + (\delta_{il} \delta_{jm} B_j \partial_l A_m - \delta_{im} \delta_{jl} B_j \partial_l A_m)$$

$$+ (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(distributivity)

$$\delta_{il} \delta_{jm} A_j \partial_l B_m - A_l \partial_l B_i + \delta_{il} \delta_{jm} B_j \partial_l A_m - B_l \partial_l A_i + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(Eq. 134)

$$\delta_{il} \delta_{jm} A_j \partial_l B_m - (A_l \partial_l) B_i + \delta_{il} \delta_{jm} B_j \partial_l A_m - (B_l \partial_l) A_i + (A_l \partial_l) B_i + (B_l \partial_l) A_i =$$

(grouping)

$$\delta_{il} \delta_{jm} A_j \partial_l B_m + \delta_{il} \delta_{jm} B_j \partial_l A_m =$$

(cancellation)

$$A_m \partial_i B_m + B_m \partial_i A_m =$$

(Eq. 134)

$\partial_i (A_m B_m) =$

(product rule)

$[\nabla (\mathbf{A} \cdot \mathbf{B})]_i$

(Eqs. 186 & 188)

Because i is a free index, the identity is proved for all components.

• $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$:

$$\begin{aligned}
\nabla \cdot (\mathbf{A} \times \mathbf{B}) &= \partial_i [\mathbf{A} \times \mathbf{B}]_i && \text{(Eq. 188)} \\
&= \partial_i (\epsilon_{ijk} A_j B_k) && \text{(Eq. 173)} \\
&= \epsilon_{ijk} \partial_i (A_j B_k) && (\partial \text{ not acting on } \epsilon) \\
&= \epsilon_{ijk} (B_k \partial_i A_j + A_j \partial_i B_k) && \text{(product rule)} \\
&= \epsilon_{ijk} B_k \partial_i A_j + \epsilon_{ijk} A_j \partial_i B_k && \text{(distributivity)} \\
&= \epsilon_{kij} B_k \partial_i A_j - \epsilon_{jik} A_j \partial_i B_k && \text{(Eq. 142)} \\
&= B_k (\epsilon_{kij} \partial_i A_j) - A_j (\epsilon_{jik} \partial_i B_k) && \text{(commutativity \& grouping)} \\
&= B_k [\nabla \times \mathbf{A}]_k - A_j [\nabla \times \mathbf{B}]_j && \text{(Eq. 192)} \\
&= \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}) && \text{(Eq. 172)}
\end{aligned}$$

• $\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla) \mathbf{A} + (\nabla \cdot \mathbf{B}) \mathbf{A} - (\nabla \cdot \mathbf{A}) \mathbf{B} - (\mathbf{A} \cdot \nabla) \mathbf{B}$:

$$\begin{aligned}
[\nabla \times (\mathbf{A} \times \mathbf{B})]_i &= \epsilon_{ijk} \partial_j [\mathbf{A} \times \mathbf{B}]_k && \text{(Eq. 192)} \\
&= \epsilon_{ijk} \partial_j (\epsilon_{klm} A_l B_m) && \text{(Eq. 173)} \\
&= \epsilon_{ijk} \epsilon_{klm} \partial_j (A_l B_m) && (\partial \text{ not acting on } \epsilon) \\
&= \epsilon_{ijk} \epsilon_{klm} (B_m \partial_j A_l + A_l \partial_j B_m) && \text{(product rule)} \\
&= \epsilon_{ijk} \epsilon_{lmk} (B_m \partial_j A_l + A_l \partial_j B_m) && \text{(Eq. 142)}
\end{aligned}$$

$$= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})(B_m\partial_j A_l + A_l\partial_j B_m) \qquad \text{(Eq. 151)}$$

$$= \delta_{il}\delta_{jm}B_m\partial_j A_l + \delta_{il}\delta_{jm}A_l\partial_j B_m -$$

$$\delta_{im}\delta_{jl}B_m\partial_j A_l - \delta_{im}\delta_{jl}A_l\partial_j B_m \qquad \text{(distributivity)}$$

$$= B_m\partial_m A_i + A_i\partial_m B_m - B_i\partial_j A_j - A_j\partial_j B_i \qquad \text{(Eq. 134)}$$

$$= (B_m\partial_m) A_i + (\partial_m B_m) A_i - (\partial_j A_j) B_i - (A_j\partial_j) B_i \qquad \text{(grouping)}$$

$$= [(\mathbf{B}\cdot\nabla)\mathbf{A}]_i + [(\nabla\cdot\mathbf{B})\mathbf{A}]_i -$$

$$[(\nabla\cdot\mathbf{A})\mathbf{B}]_i - [(\mathbf{A}\cdot\nabla)\mathbf{B}]_i \qquad \text{(Eqs. 196 \& 188)}$$

$$= [(\mathbf{B}\cdot\nabla)\mathbf{A} + (\nabla\cdot\mathbf{B})\mathbf{A} - (\nabla\cdot\mathbf{A})\mathbf{B} - (\mathbf{A}\cdot\nabla)\mathbf{B}]_i \qquad \text{(Eq. 66)}$$

Because i is a free index, the identity is proved for all components.

- $(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = \begin{vmatrix} \mathbf{A}\cdot\mathbf{C} & \mathbf{A}\cdot\mathbf{D} \\ \mathbf{B}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{D} \end{vmatrix}$:

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = [\mathbf{A}\times\mathbf{B}]_i [\mathbf{C}\times\mathbf{D}]_i \qquad \text{(Eq. 172)}$$

$$= \epsilon_{ijk}A_jB_k\epsilon_{ilm}C_lD_m \qquad \text{(Eq. 173)}$$

$$= \epsilon_{ijk}\epsilon_{ilm}A_jB_kC_lD_m \qquad \text{(commutativity)}$$

$$= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) A_jB_kC_lD_m \qquad \text{(Eqs. 142 \& 151)}$$

$$= \delta_{jl}\delta_{km}A_jB_kC_lD_m - \delta_{jm}\delta_{kl}A_jB_kC_lD_m \qquad \text{(distributivity)}$$

$$= (\delta_{jl}A_jC_l)(\delta_{km}B_kD_m) - (\delta_{jm}A_jD_m)(\delta_{kl}B_kC_l) \qquad \text{(commutativity)}$$

$$= (A_lC_l)(B_mD_m) - (A_mD_m)(B_lC_l) \qquad \text{(Eq. 134)}$$

$$= (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C}) \qquad \text{(Eq. 172)}$$

$$= \begin{vmatrix} \mathbf{A}\cdot\mathbf{C} & \mathbf{A}\cdot\mathbf{D} \\ \mathbf{B}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{D} \end{vmatrix} \qquad \text{(definition)}$$

- $(\mathbf{A}\times\mathbf{B})\times(\mathbf{C}\times\mathbf{D}) = [\mathbf{D}\cdot(\mathbf{A}\times\mathbf{B})]\,\mathbf{C} - [\mathbf{C}\cdot(\mathbf{A}\times\mathbf{B})]\,\mathbf{D}$:

$$
\begin{aligned}
\left[(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D})\right]_i &= \epsilon_{ijk} \left[\mathbf{A} \times \mathbf{B}\right]_j \left[\mathbf{C} \times \mathbf{D}\right]_k && \text{(Eq. 173)} \\
&= \epsilon_{ijk} \epsilon_{jmn} A_m B_n \epsilon_{kpq} C_p D_q && \text{(Eq. 173)} \\
&= \epsilon_{ijk} \epsilon_{kpq} \epsilon_{jmn} A_m B_n C_p D_q && \text{(commutativity)} \\
&= \epsilon_{ijk} \epsilon_{pqk} \epsilon_{jmn} A_m B_n C_p D_q && \text{(Eq. 142)} \\
&= \left(\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}\right) \epsilon_{jmn} A_m B_n C_p D_q && \text{(Eq. 151)} \\
&= \left(\delta_{ip}\delta_{jq}\epsilon_{jmn} - \delta_{iq}\delta_{jp}\epsilon_{jmn}\right) A_m B_n C_p D_q && \text{(distributivity)} \\
&= \left(\delta_{ip}\epsilon_{qmn} - \delta_{iq}\epsilon_{pmn}\right) A_m B_n C_p D_q && \text{(Eq. 134)} \\
&= \delta_{ip}\epsilon_{qmn} A_m B_n C_p D_q - \delta_{iq}\epsilon_{pmn} A_m B_n C_p D_q && \text{(distributivity)} \\
&= \epsilon_{qmn} A_m B_n C_i D_q - \epsilon_{pmn} A_m B_n C_p D_i && \text{(Eq. 134)} \\
&= \epsilon_{qmn} D_q A_m B_n C_i - \epsilon_{pmn} C_p A_m B_n D_i && \text{(commutativity)} \\
&= \left(\epsilon_{qmn} D_q A_m B_n\right) C_i - \left(\epsilon_{pmn} C_p A_m B_n\right) D_i && \text{(grouping)} \\
&= \left[\mathbf{D} \cdot (\mathbf{A} \times \mathbf{B})\right] C_i - \left[\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})\right] D_i && \text{(Eq. 174)} \\
&= \left[\left[\mathbf{D} \cdot (\mathbf{A} \times \mathbf{B})\right] \mathbf{C}\right]_i - \left[\left[\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})\right] \mathbf{D}\right]_i && \text{(index definition)} \\
&= \left[\left[\mathbf{D} \cdot (\mathbf{A} \times \mathbf{B})\right] \mathbf{C} - \left[\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})\right] \mathbf{D}\right]_i && \text{(Eq. 66)}
\end{aligned}
$$

Because i is a free index, the identity is proved for all components.

5.7 Exercises

5.1 Write, in tensor notation, the mathematical expression for the trace, determinant and inverse of an $n \times n$ matrix.

5.2 Repeat the previous exercise for the multiplication of a matrix by a vector and the multiplication of two $n \times n$ matrices.

5.3 Define mathematically the dot and cross product operations of two vectors using tensor notation.

5.4 Repeat the previous exercise for the scalar triple product and vector triple product operations of three vectors.

5.5 Define mathematically, using tensor notation, the three scalar invariants of a rank-2 tensor: I, II and III.

5.6 Express the three scalar invariants I, II and III in terms of the other three invariants I_1, I_2 and I_3 and vice versa.

5.7 Explain why the three invariants I_1, I_2 and I_3 are scalars using in your argument the fact that the three main invariants I, II and III are traces?

5.8 Gather six terms from the Index about the scalar invariants of tensors.

5.9 Justify, giving a detailed explanation, the following statement: "If a rank-2 tensor is invertible in a particular coordinate system it is invertible in all other coordinate systems, and if it is singular in a particular coordinate system it is singular in all other coordinate systems". Use in your explanation the fact that the determinant is invariant under admissible coordinate transformations.

5.10 What are the ten joint invariants between two rank-2 tensors?

5.11 Provide a concise mathematical definition of the nabla differential operator ∇ in Cartesian coordinate systems using tensor notation.

5.12 What is the rank and variance type of the gradient of a differentiable scalar field in general curvilinear coordinate systems?

5.13 State, in tensor notation, the mathematical expression for the gradient of a differentiable scalar field in a Cartesian system.

5.14 What is the gradient of the following scalar functions of position f, g and h where x_1, x_2 and x_3 are the Cartesian coordinates and a, b and c are constants?

$$f = 1.3x_1 - 2.6ex_2 + 19.8x_3 \qquad g = ax_3 + be^{x_2} \qquad h = a(x_1)^3 - \sin x_3 + c(x_3)^2$$

5.15 State, in tensor notation, the mathematical expression for the gradient of a differentiable vector field in a Cartesian system.

5.16 What is the gradient of the following vector where x_1, x_2 and x_3 are the Cartesian coordinates?

$$\mathbf{V} = (2x_1 - 1.2x_2,\ x_1 + x_3,\ x_2x_3)$$

What is the rank of this gradient?

5.17 Explain, in detail, why the divergence of a vector is invariant.

5.18 What is the rank of the divergence of a rank-n ($n > 0$) tensor and why?

5.19 State, using vector and tensor notations, the mathematical definition of the divergence operation of a vector in a Cartesian coordinate system.

5.20 Discuss in detail the following statement: "The divergence of a vector is a gradient operation followed by a contraction". How this is related to the trace of a rank-2 tensor?

5.21 Write down the mathematical expression of the two forms of the divergence of a rank-2 tensor.

5.22 How many forms do we have for the divergence of a rank-n ($n > 0$) tensor and why? Assume in your answer that the divergence operation can be conducted with respect to any one of the tensor indices.

5.23 Find the divergence of the following vectors **U** and **V** where x_1, x_2 and x_3 are the Cartesian coordinates:

$$\mathbf{U} = \left(9.3x_1,\ 6.3\cos x_2,\ 3.6x_1 e^{-1.2x_3}\right) \qquad \mathbf{V} = \left(x_2 \sin x_1,\ 5(x_2)^3,\ 16.3x_3\right)$$

5.24 State, in tensor notation, the mathematical expression for the curl of a vector and of a rank-2 tensor assuming a Cartesian coordinate system.

5.25 By using the Index, make a list of terms and notations related to matrix algebra which are used in this chapter.

5.26 Define, in tensor notation, the Laplacian operator acting on a differentiable scalar field in a Cartesian coordinate system.

5.27 Is the Laplacian a scalar or a vector operator?

5.28 What is the meaning of the Laplacian operator acting on a differentiable vector field?

5.29 What is the rank of a rank-n tensor acted upon by the Laplacian operator?

5.30 Define mathematically the following operators assuming a Cartesian coordinate system:

$$\mathbf{A} \cdot \nabla \qquad \mathbf{A} \times \nabla \tag{239}$$

What is the rank of each one of these operators?

5.31 Make a general statement about how differentiation of tensors affects their rank discussing in detail from this perspective the gradient and divergence operations.

5.32 State the mathematical expressions for the following operators and operations assuming a cylindrical coordinate system: nabla operator, Laplacian operator, gradient of a scalar, divergence of a vector, and curl of a vector.

5.33 Explain how the expressions for the operators and operations in the previous exercise can be obtained for the plane polar coordinate system from the expressions of the cylindrical system.

5.34 State the mathematical expressions for the following operators and operations assuming a spherical coordinate system: nabla operator, Laplacian operator, gradient of a scalar, divergence of a vector, and curl of a vector.

5.35 Repeat the previous exercise for the general orthogonal coordinate system.

5.36 Express, in tensor notation, the mathematical condition for a vector field to be solenoidal.

5.37 Express, in tensor notation, the mathematical condition for a vector field to be irrotational.

5.38 Express, in tensor notation, the divergence theorem for a differentiable vector field explaining all the symbols involved. Repeat the exercise for a differentiable tensor field of an arbitrary rank (> 0).

5.39 Express, in tensor notation, Stokes theorem for a differentiable vector field explaining all the symbols involved. Repeat the exercise for a differentiable tensor field of an arbitrary rank (> 0).

5.40 Express the following identities in tensor notation:

$$\begin{aligned}
\nabla \cdot \mathbf{r} &= n \\
\nabla (\mathbf{a} \cdot \mathbf{r}) &= \mathbf{a} \\
\nabla \cdot (\nabla \times \mathbf{A}) &= 0 \\
\nabla (fh) &= f\nabla h + h\nabla f
\end{aligned}$$

$$\begin{aligned}
\nabla \times (f\mathbf{A}) &= f\nabla \times \mathbf{A} + \nabla f \times \mathbf{A} \\
\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= \mathbf{B}\,(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}\,(\mathbf{A} \cdot \mathbf{B}) \\
\nabla \times (\nabla \times \mathbf{A}) &= \nabla\,(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \\
\nabla \cdot (\mathbf{A} \times \mathbf{B}) &= \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}) \\
(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= \begin{vmatrix} \mathbf{A} \cdot \mathbf{C} & \mathbf{A} \cdot \mathbf{D} \\ \mathbf{B} \cdot \mathbf{C} & \mathbf{B} \cdot \mathbf{D} \end{vmatrix}
\end{aligned}$$

5.41 Prove the following identities using the language and techniques of tensor calculus:

$$\begin{aligned}
\nabla \times \mathbf{r} &= \mathbf{0} \\
\nabla \cdot (\nabla f) &= \nabla^2 f \\
\nabla \times (\nabla f) &= \mathbf{0} \\
\nabla \cdot (f\mathbf{A}) &= f\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla f \\
\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) &= \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) \\
\mathbf{A} \times (\nabla \times \mathbf{B}) &= (\nabla \mathbf{B}) \cdot \mathbf{A} - \mathbf{A} \cdot \nabla \mathbf{B} \\
\nabla\,(\mathbf{A} \cdot \mathbf{B}) &= \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\,\mathbf{B} + (\mathbf{B} \cdot \nabla)\,\mathbf{A} \\
\nabla \times (\mathbf{A} \times \mathbf{B}) &= (\mathbf{B} \cdot \nabla)\,\mathbf{A} + (\nabla \cdot \mathbf{B})\,\mathbf{A} - (\nabla \cdot \mathbf{A})\,\mathbf{B} - (\mathbf{A} \cdot \nabla)\,\mathbf{B} \\
(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) &= [\mathbf{D} \cdot (\mathbf{A} \times \mathbf{B})]\,\mathbf{C} - [\mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})]\,\mathbf{D}
\end{aligned}$$

Chapter 6

Metric Tensor

The subject of the present chapter is the metric tensor which is one of the most important special tensors, if not the most important of all, in tensor calculus. Its versatile usage and functionalities permeate the whole discipline of tensor calculus.

The metric tensor is a **rank-2** tensor which may also be called the **fundamental tensor**. The main purpose of the metric tensor is to generalize the concept of **distance** to general curvilinear coordinate frames and maintain the **invariance** of distance in different coordinate systems.

In orthonormal Cartesian coordinate systems the **distance element** squared, $(ds)^2$, between two infinitesimally neighboring points in space, one with coordinates x_i and the other with coordinates $x_i + dx_i$, is given by:

$$(ds)^2 = dx_i dx_i = \delta_{ij} dx_i dx_j \tag{240}$$

This definition of distance is the key to introducing a rank-2 tensor, g_{ij}, called the **metric tensor** which, for a general coordinate system, is defined by:

$$(ds)^2 = g_{ij} dx^i dx^j \tag{241}$$

The above defined metric tensor is of **covariant** type. The metric tensor has also a **contravariant** type which is usually notated with g^{ij}.

The **components** of the covariant and contravariant metric tensor are given by:

$$g_{ij} = \mathbf{E}_i \cdot \mathbf{E}_j \qquad\qquad g^{ij} = \mathbf{E}^i \cdot \mathbf{E}^j \tag{242}$$

where the indexed **E** are the covariant and contravariant **basis vectors** as defined in § 2.6.1.

The metric tensor has also a **mixed** type which is given by:

$$g^i{}_j = \mathbf{E}^i \cdot \mathbf{E}_j = \delta^i{}_j \qquad\qquad g_i{}^j = \mathbf{E}_i \cdot \mathbf{E}^j = \delta_i{}^j \tag{243}$$

and hence it is the same as the **unity tensor**.

For a coordinate system in which the metric tensor can be cast in a **diagonal** form where the diagonal elements are ± 1 the metric is called **flat**. For Cartesian coordinate systems, which are orthonormal flat-space systems, we have:

$$g^{ij} = \delta^{ij} = g_{ij} = \delta_{ij} \tag{244}$$

The metric tensor is **symmetric** in its two indices, that is:

$$g_{ij} = g_{ji} \qquad\qquad g^{ij} = g^{ji} \tag{245}$$

This can be easily explained by the commutativity of the dot product of vectors in reference to the above equations involving the dot product of the basis vectors.

The contravariant metric tensor is used for **raising** covariant indices of covariant and mixed tensors, e.g.

$$A^i = g^{ik} A_k \qquad\qquad A^{ij} = g^{ik} A_k{}^j \tag{246}$$

Similarly, the covariant metric tensor is used for **lowering** contravariant indices of contravariant and mixed tensors, e.g.

$$A_i = g_{ik}A^k \qquad\qquad A_{ij} = g_{ik}A^k{}_j \tag{247}$$

In these raising and lowering operations the metric tensor acts, like a Kronecker delta, as an index **replacement** operator as well as shifting the position of the index.

Because it is possible to shift the index position of a tensor by using the covariant and contravariant types of the metric tensor, a given tensor **can be cast** into a covariant or a contravariant form, as well as a mixed form in the case of tensors of rank > 1. However, it should be emphasized that the **order** of the indices must be respected in this process, because two tensors with the same indicial structure but with different indicial order are not equal in general, as stated before. For example:

$$A^i{}_j = g_{jk}A^{ik} \neq A_j{}^i = g_{jk}A^{ki} \tag{248}$$

Some authors insert **dots** (e.g. $A^i_{\cdot j}$ and $A_j^{\cdot i}$) to remove any ambiguity about the order of the indices.

The covariant and contravariant metric tensors are **inverses** of each other, that is:

$$[g_{ij}] = \left[g^{ij}\right]^{-1} \qquad\qquad \left[g^{ij}\right] = [g_{ij}]^{-1} \tag{249}$$

Hence:

$$g^{ik}g_{kj} = \delta^i{}_j \qquad\qquad g_{ik}g^{kj} = \delta_i{}^j \tag{250}$$

It is common to reserve the term **metric tensor** to the covariant form and call the contravariant form, which is its inverse, the **associate** or **conjugate** or **reciprocal metric tensor**. As a tensor, the metric has a significance regardless of any coordinate system although it requires a coordinate system to be represented in a specific form. For **orthog-**

onal coordinate systems the metric tensor is **diagonal**, i.e. $g_{ij} = g^{ij} = 0$ for $i \neq j$.

As indicated before, for orthonormal **Cartesian** coordinate systems in a 3D space, the metric tensor is given in its covariant and contravariant forms by the 3×3 unit matrix, that is:

$$[g_{ij}] = [\delta_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \left[\delta^{ij}\right] = \left[g^{ij}\right] \tag{251}$$

For **cylindrical** coordinate systems with coordinates (ρ, ϕ, z), the metric tensor is given in its covariant and contravariant forms by:

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \left[g^{ij}\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho^2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{252}$$

while for **spherical** coordinate systems with coordinates (r, θ, ϕ), the metric tensor is given in its covariant and contravariant forms by:

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{bmatrix} \qquad \left[g^{ij}\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2\theta} \end{bmatrix} \tag{253}$$

As seen, all these metric tensors are **diagonal** since all these coordinate systems are orthogonal. We also notice that all the corresponding diagonal elements of the covariant and contravariant types are **reciprocals** of each other. This can be easily explained by the fact that these two types are inverses of each other, plus the fact that the inverse of an invertible diagonal matrix is a diagonal matrix obtained by taking the reciprocal of the corresponding diagonal elements of the original matrix, as stated in § Inverse of Matrix.

6.1 Exercises

6.1 Describe in details, using mathematical tensor language when necessary, the metric tensor discussing its rank, purpose, designations, variance types, symmetry, its role in the definition of distance, and its relation to the covariant and contravariant basis vectors.

6.2 What is the relation between the covariant and contravariant types of the metric tensor? Express this relation mathematically. Also define mathematically the mixed type metric tensor.

6.3 Correct, if necessary, the following equations:

$$g^i{}_j = \delta^i{}_i \qquad g^{ij} = \mathbf{E}^i \cdot \mathbf{E}_j \qquad (ds) = g_{ij} dx^i dx^j \qquad g_{ij} = \mathbf{E}_i \cdot \mathbf{E}_j \qquad \mathbf{E}^i \cdot \mathbf{E}_j = \delta^j{}_i$$

6.4 What "flat metric" means? Give an example of a coordinate system with a flat metric.

6.5 Describe the index-shifting (raising/lowering) operators and their relation to the metric tensor. How these operators facilitate the transformation between the covariant, contravariant and mixed types of a given tensor?

6.6 Find from the Index all the names and labels of the metric tensor in its covariant, contravariant and mixed types.

6.7 What is wrong with the following equations?

$$C_i = g^{ij} C_j \qquad D_i = g^{ij} D^j \qquad A^i = \delta_{ij} A_j$$

Make the necessary corrections considering all the possibilities in each case.

6.8 Is it necessary to keep the order of the indices which are shifted by the index-shifting operators and why?

6.9 How and why dots may be inserted to avoid confusion about the order of the indices following an index-shifting operation?

6.10 Express, mathematically, the fact that the contravariant and covariant metric tensors are inverses of each other.

6.11 Collect from the Index a list of operators and operations which have particular links to the metric tensor.

6.12 Correct, if necessary, the following statement: "The term metric tensor is usually used to label the covariant form of the metric, while the contravariant form of the metric is called the conjugate or associate or reciprocal metric tensor".

6.13 Write, in matrix form, the covariant and contravariant types of the metric tensor of the Cartesian, cylindrical and spherical coordinate systems of a 3D flat space.

6.14 Regarding the previous question, what do you notice about the corresponding diagonal elements of the covariant and contravariant types of the metric tensor in these systems? Does this relate to the fact that these types are inverses of each other?

Chapter 7

Covariant Differentiation

The focus of this chapter is the operation of covariant differentiation of tensors which, in a sense, is a **generalization** of the ordinary differentiation. The ordinary derivative of a tensor is not a tensor in general. The objective of covariant differentiation is to ensure the **invariance** of derivative (i.e. being a tensor) in general coordinate systems, and this results in applying more sophisticated rules using Christoffel symbols where different differentiation rules for covariant and contravariant indices apply. The resulting covariant derivative is a **tensor** which is one rank higher than the differentiated tensor.

The **Christoffel symbol** of the second kind is defined by:

$$\left\{ {k \atop ij} \right\} = \frac{g^{kl}}{2} \left(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l} \right) \tag{254}$$

where the indexed g is the metric tensor in its contravariant and covariant forms with implied summation over l. It is noteworthy that Christoffel symbols are **not tensors**. The Christoffel symbols of the second kind are **symmetric** in their two lower indices, that is:

$$\left\{ {k \atop ij} \right\} = \left\{ {k \atop ji} \right\} \tag{255}$$

For **Cartesian** coordinate systems, the Christoffel symbols are zero for all values of the indices. For **cylindrical** coordinate systems, marked with the coordinates (ρ, ϕ, z), the

Christoffel symbols are zero for all values of the indices except:

$$\left\{ {1 \atop 22} \right\} = -\rho \tag{256}$$

$$\left\{ {2 \atop 12} \right\} = \left\{ {2 \atop 21} \right\} = \frac{1}{\rho} \tag{257}$$

where $(1, 2, 3)$ stand for (ρ, ϕ, z).

For **spherical** coordinate systems, marked with the coordinates (r, θ, ϕ), the Christoffel symbols are zero for all values of the indices except:

$$\left\{ {1 \atop 22} \right\} = -r \tag{258}$$

$$\left\{ {1 \atop 33} \right\} = -r \sin^2 \theta \tag{259}$$

$$\left\{ {2 \atop 12} \right\} = \left\{ {2 \atop 21} \right\} = \frac{1}{r} \tag{260}$$

$$\left\{ {2 \atop 33} \right\} = -\sin \theta \cos \theta \tag{261}$$

$$\left\{ {3 \atop 13} \right\} = \left\{ {3 \atop 31} \right\} = \frac{1}{r} \tag{262}$$

$$\left\{ {3 \atop 23} \right\} = \left\{ {3 \atop 32} \right\} = \cot \theta \tag{263}$$

where $(1, 2, 3)$ stand for (r, θ, ϕ).

For a differentiable **scalar** f, the covariant derivative is the **same** as the ordinary partial derivative, that is:

$$f_{;i} = f_{,i} = \partial_i f \tag{264}$$

This is justified by the fact that the covariant derivative is different from the ordinary partial derivative because the basis vectors in general coordinate systems are dependent on their spatial position, and since a scalar is independent of the basis vectors the covariant and partial derivatives are identical.

For a differentiable **vector A**, the covariant derivative of the covariant and contravariant forms of the vector is given by:

$$A_{j;i} = \partial_i A_j - \left\{ {k \atop ji} \right\} A_k \qquad \text{(covariant)} \qquad (265)$$

$$A^j{}_{;i} = \partial_i A^j + \left\{ {j \atop ki} \right\} A^k \qquad \text{(contravariant)} \qquad (266)$$

For a differentiable **rank-2** tensor **A**, the covariant derivative of the covariant, contravariant and mixed forms of the tensor is given by:

$$A_{jk;i} = \partial_i A_{jk} - \left\{ {l \atop ji} \right\} A_{lk} - \left\{ {l \atop ki} \right\} A_{jl} \qquad \text{(covariant)} \qquad (267)$$

$$A^{jk}{}_{;i} = \partial_i A^{jk} + \left\{ {j \atop li} \right\} A^{lk} + \left\{ {k \atop li} \right\} A^{jl} \qquad \text{(contravariant)} \qquad (268)$$

$$A^k_{j;i} = \partial_i A^k_j + \left\{ {k \atop li} \right\} A^l_j - \left\{ {l \atop ji} \right\} A^k_l \qquad \text{(mixed)} \qquad (269)$$

More generally, for a differentiable **rank-n tensor A**, the covariant derivative is given by:

$$\begin{aligned} A^{ij\ldots k}_{lm\ldots p;q} = {} & \partial_q A^{ij\ldots k}_{lm\ldots p} + \left\{ {i \atop aq} \right\} A^{aj\ldots k}_{lm\ldots p} + \left\{ {j \atop aq} \right\} A^{ia\ldots k}_{lm\ldots p} + \cdots + \left\{ {k \atop aq} \right\} A^{ij\ldots a}_{lm\ldots p} \\ & - \left\{ {a \atop lq} \right\} A^{ij\ldots k}_{am\ldots p} - \left\{ {a \atop mq} \right\} A^{ij\ldots k}_{la\ldots p} - \cdots - \left\{ {a \atop pq} \right\} A^{ij\ldots k}_{lm\ldots a} \end{aligned} \qquad (270)$$

From the last equations, a **pattern** for the covariant differentiation operation emerges, that is it starts with an ordinary partial derivative term, then for each tensor index an extra Christoffel symbol term is added, positive for superscripts and negative for subscripts, where the differentiation index is the second of the lower indices in the Christoffel symbol.

Since the Christoffel symbols are identically zero in **Cartesian** coordinate systems, the covariant derivative is the **same** as the ordinary partial derivative for all tensor ranks. Another important fact about covariant differentiation is that the covariant derivative of the **metric tensor** in its covariant, contravariant and mixed forms is **zero** in all coordinate systems and hence it is treated like a constant in covariant differentiation.

Several **rules** of ordinary differentiation **similarly apply** to covariant differentiation. For example, covariant differentiation is a **linear** operation with respect to algebraic sums of tensor terms, that is:

$$\partial_{;i}\left(a\mathbf{A} \pm b\mathbf{B}\right) = a\,\partial_{;i}\mathbf{A} \pm b\,\partial_{;i}\mathbf{B} \tag{271}$$

where a and b are scalar constants and $\mathbf{A}$ and $\mathbf{B}$ are differentiable tensor fields.

The **product rule** of ordinary differentiation also applies to covariant differentiation of tensor multiplication, that is:

$$\partial_{;i}\left(\mathbf{AB}\right) = \left(\partial_{;i}\mathbf{A}\right)\mathbf{B} + \mathbf{A}\partial_{;i}\mathbf{B} \tag{272}$$

However, as seen in this equation, the order of the tensors should be observed since tensor multiplication, unlike ordinary algebraic multiplication, is not commutative. The product rule is also valid for the **inner product** of tensors because the inner product is an outer product operation followed by a contraction of indices, and covariant differentiation and contraction of indices do commute.

Since the covariant derivative of the metric tensor is identically zero, as stated above, the covariant derivative operator **bypasses** the index raising/lowering operator, that is:

$$\partial_{;m}\left(g_{ij}A^{j}\right) = g_{ij}\partial_{;m}A^{j} \tag{273}$$

and hence the metric tensor behaves like a **constant** with respect to the covariant differential operator.

A principal **difference** between the partial differentiation and the covariant differentiation is that for successive differential operations with respect to different indices the partial derivative operators do **commute** with each other, assuming certain continuity conditions, but the covariant differential operators **do not commute**, that is:

$$\partial_i \partial_j = \partial_j \partial_i \qquad \text{but} \qquad \partial_{;i} \partial_{;j} \neq \partial_{;j} \partial_{;i} \tag{274}$$

Higher order covariant derivatives are **similarly defined** as derivatives of derivatives; however the order of differentiation, in the case of differentiating with respect to different indices, should be respected, as explained above.

7.1 Exercises

7.1 Is the ordinary derivative of a tensor necessarily a tensor or not? Can the ordinary derivative of a tensor be a tensor? If so, give an example.

7.2 Explain the purpose of the covariant derivative and how it is related to the invariance property of tensors.

7.3 Is the covariant derivative of a tensor necessarily a tensor? If so, what is the rank of the covariant derivative of a rank-n tensor?

7.4 How is the Christoffel symbol of the second kind symbolized? Describe the arrangement of its indices.

7.5 State the mathematical definition of the Christoffel symbol of the second kind in terms of the metric tensor defining all the symbols involved.

7.6 The Christoffel symbols of the second kind are symmetric in which of their indices?

7.7 Why the Christoffel symbols of the second kind are identically zero in the Cartesian coordinate systems? Use in your explanation the mathematical definition of the Christoffel symbols.

7.8 Give the Christoffel symbols of the second kind for the cylindrical and spherical coordinate systems explaining the meaning of the indices used.

7.9 From the Index, collect five terms used in the definition and description of the Christoffel symbols of the second kind.

7.10 What is the meaning, within the context of tensor differentiation, of the comma "," and semicolon ";" when used as subscripts preceding a tensor index?

7.11 Why the covariant derivative of a differentiable scalar is the same as the ordinary partial derivative and how is this related to the basis vectors of coordinate systems?

7.12 Differentiate the following tensors covariantly:

$$A^s \qquad B_t \qquad C_i^j \qquad D_{pq} \qquad E^{mn} \qquad A_{ij...k}^{lm...p}$$

7.13 Explain the mathematical pattern followed in the operation of covariant differentiation of tensors. Does this pattern also apply to rank-0 tensors?

7.14 The covariant derivative in Cartesian coordinate systems is the same as the ordinary partial derivative for all tensor ranks. Explain why.

7.15 What is the covariant derivative of the covariant and contravariant forms of the metric tensor for an arbitrary type of coordinate system? How is this related to the fact that the covariant derivative operator bypasses the index-shifting operator?

7.16 Which rules of ordinary differentiation apply equally to covariant differentiation and which do not? Make mathematical statements about all these rules with sufficient explanation of the symbols and operations involved.

7.17 Make corrections, where necessary, in the following equations explaining in each case why the equation should or should not be amended:

$$
\begin{aligned}
(\mathbf{C} \pm \mathbf{D})_{;i} &= \partial_{;i}\mathbf{C} \pm \partial_{;i}\mathbf{D} \\
\partial_{;i}(\mathbf{AB}) &= \mathbf{B}(\partial_{;i}\mathbf{A}) + \mathbf{A}\partial_{;i}\mathbf{B} \\
\left(g_{ij}A^j\right)_{;m} &= g_{ij}\left(A^j\right)_{;m} \\
\partial_{;i}\partial_{;j} &= \partial_{;j}\partial_{;i} \\
\partial_i\partial_j &= \partial_j\partial_i
\end{aligned}
$$

7.18 How do you define the second and higher order covariant derivatives of tensors? Do these derivatives follow the same rules as the ordinary partial derivatives of the same order in the case of different differentiation indices?

7.19 From the Index, find all the terms that refer to symbols used in the notation of ordinary partial derivatives and covariant derivatives.

References

G.B. Arfken; H.J. Weber; F.E. Harris. *Mathematical Methods for Physicists A Comprehensive Guide*. Elsevier Academic Press, seventh edition, 2013.

R.B. Bird; R.C. Armstrong; O. Hassager. *Dynamics of Polymeric Liquids*, volume 1. John Wiley & Sons, second edition, 1987.

R.B. Bird; W.E. Stewart; E.N. Lightfoot. *Transport Phenomena*. John Wiley & Sons, second edition, 2002.

M.L. Boas. *Mathematical Methods in the Physical Sciences*. John Wiley & Sons Inc., third edition, 2006.

C.F. Chan Man Fong; D. De Kee; P.N. Kaloni. *Advanced Mathematics for Engineering and Science*. World Scientific Publishing Co. Pte. Ltd., first edition, 2003.

T.L. Chow. *Mathematical Methods for Physicists: A concise introduction*. Cambridge University Press, first edition, 2003.

J.H. Heinbockel. *Introduction to Tensor Calculus and Continuum Mechanics*. 1996.

D.C. Kay. *Schaum's Outline of Theory and Problems of Tensor Calculus*. McGraw-Hill, first edition, 1988.

K.F. Riley; M.P. Hobson; S.J. Bence. *Mathematical Methods for Physics and Engineering*. Cambridge University Press, third edition, 2006.

D. Zwillinger, editor. *CRC Standard Mathematical Tables and Formulae*. CRC Press, 32nd edition, 2012.

Index

About the Book

This book is prepared from personal notes and tutorials about tensor calculus at an introductory level. The language and method used in presenting the ideas and techniques of tensor calculus make it very suitable for learning this subject by the beginners who have not been exposed previously to this elegant branch of mathematics. Considerable efforts have been made to reduce the dependency on foreign texts by summarizing the main concepts needed to make the book self-contained. The book also contains a number of good-quality graphic illustrations to aid the readers and students in their effort to visualize the ideas and understand the abstract concepts. Furthermore, illustrative techniques, such as highlighting key terms by boldface fonts, have been employed. The book also contains sets of clearly explained exercises which cover most of the given materials. These exercises are designed to provide thorough revisions of the supplied materials and hence they make an essential component of the book and its learning objectives. The book is also furnished with a rather detailed index and populated with cross references, which are hyperlinked for the ebook users, to facilitate connecting related subjects and ideas.

About the Author

The author of the book possesses a diverse academic and research background. He holds a BEng in electronics engineering, a BSc in physics, a PhD in petroleum engineering, a PhD in crystallography, and a PhD in atomic physics and astronomy. He also worked as a postdoctoral research associate for several years. He published dozens of scientific research papers in many refereed journals and produced numerous academically oriented documents which are freely available on the world wide web.