1

An Introduction to Tensor Calculus

1.1 Overall Context

We will be developing the laws of continuum physics throughout the first part of this book. We will do so in an uncommon but pedagogic way by starting with the laws that describe the discrete movement of individual atoms and then summing over the molecular dynamics. The emergent continuum laws so obtained come in the form of partial-differential equations (PDEs) that determine how fields are changing in time at each point in space based on how the fields are varying in space in the immediate neighborhood of that point. The second part of the book will treat, for the most part, mathematical techniques for analytically solving the PDEs with a heavy dose of Fourier analysis and contour-integration methods. Students learning this material from me over the years have reported that the first part where the continuum rules are established is more difficult for them compared to the second part where math problems are solved. Perhaps this is analogous to how building a toy model is more challenging than playing with the toy once it is built.

In continuum mechanics in particular, the key field representing the underlying molecular-force interactions is a tensor (the "stress tensor") and across all of continuum physics, the material properties and constitutive laws are often only describable using tensors. In short, it is impossible to learn continuum physics properly without a solid foundation in tensors and tensor calculus, and this is why we begin the book with this foundational topic. What you learn in this first chapter, especially the tensor-calculus product-rule identities of Section 1.7, will be used at every step throughout our development of the rules of continuum physics. Fortunately, tensors and tensor calculus are a natural, even effortless, extension from the concepts of vectors and vector calculus that I assume you are familiar with. This chapter reviews the various types of spatial derivatives employed in continuum physics (the gradient, divergence, and curl), while allowing these spatial derivatives to act upon tensor fields, which is assumed to be new to the reader. It is my experience that even more senior research scientists can benefit from this chapter's survey of tensor calculus in preparation for the derivations in all the chapters that follow.

Throughout the book, our focus is on analytical understanding of the physics and mathematics and this involves pencil and paper work. You need to develop confidence in pushing the symbols around the page as you handle and ultimately solve the PDEs we will be deriving. The goal is to build intuition and hands-on familiarity with the physical processes being discussed. Simulating macroscopic experiments in the real world, often performed in complicated heterogeneous bodies of matter with irregular boundaries, is called the *forward problem* and usually needs to be performed numerically because analytical solutions of the governing equations are not possible. Recording the material response at places within a body during various types of experiments and minimizing the difference between the recorded data and simulations of the data with the goal of determining the physical properties throughout the body is called the *inverse problem* and is also a numerical exercise in nearly all cases. But we will not be addressing in this book numerical aspects of the forward and inverse problems posed in macroscopic bodies. Instead we content ourselves with first developing the PDEs that control basic processes of interest across many physical-science disciplines (Part I) and then solving simplified forms of the equations in simple geometries where analytical results are possible so that your physical intuition about the physics can be developed (Part II).

1.2 Some Actors

Any physical quantity continuously distributed over the space of some region is called a *field*. Continuum physics involves the study of fields. Fields can be *scalars, vectors,* or *tensors*.

Scalar Fields: A field quantity that has no intrinsic direction is called a scalar field. Examples include temperature, pressure, and various types of densities. In the nomenclature of tensors, a scalar can be called a zeroth-order tensor.

Vector Fields: A field quantity that has a direction associated with it is called a vector field. Examples include electric fields, fluid velocity, and gravitational acceleration. Vector fields are represented at each point in space by an arrow whose length denotes the amplitude of the vector field at that point. In the nomenclature of tensors, a vector can be called a first-order tensor. Vector fields as depicted in Fig. 1.1 can be written analytically in different ways:

$$\begin{aligned} \boldsymbol{r} &\,\widehat{=}\, \text{position vector used to identify points in space} \\ &= x_1\hat{\boldsymbol{x}}_1 + x_2\hat{\boldsymbol{x}}_2 + x_3\hat{\boldsymbol{x}}_3 = x\,\hat{\boldsymbol{x}} + y\,\hat{\boldsymbol{y}} + z\,\hat{\boldsymbol{z}} \\ &= (x_1, x_2, x_3) = x_i\hat{\boldsymbol{x}}_i \quad \text{(summation over repeated indices)}, \end{aligned} \tag{1.1}$$

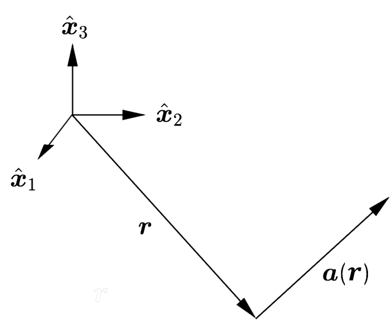

Figure 1.1 Points in space denoted by the vector $\boldsymbol{r}$ and a vector field $\boldsymbol{a}(\boldsymbol{r})$ at each point $\boldsymbol{r}$.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{r}) &= \boldsymbol{a}(x_1, x_2, x_3) \mathrel{\widehat{=}} \text{vector field defined at each point } \boldsymbol{r}\\ &= a_1\hat{\boldsymbol{x}}_1 + a_2\hat{\boldsymbol{x}}_2 + a_3\hat{\boldsymbol{x}}_3 = a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}} \\ &= (a_1, a_2, a_3) = a_i\hat{\boldsymbol{x}}_i.\end{aligned} \tag{1.2}$$

The caret symbol ˆ placed above a vector means that vector is unitless and has an amplitude of 1, that is, $\hat{\boldsymbol{a}} \mathrel{\widehat{=}} \boldsymbol{a}/|\boldsymbol{a}|$, where $|\boldsymbol{a}| = \sqrt{(a_1^2 + a_2^2 + a_3^2)}$ denotes the amplitude of vector $\boldsymbol{a}$.

IMPORTANT: Whenever an index appears twice in an expression, you always sum over that index. The index that is summed over is sometimes called a *dummy index* because the index does not survive the summation and could be given any name. For example, we have $a_ib_i = a_jb_j = a_nb_n = \sum_{n=1}^{3} a_nb_n = \sum_{j=1}^{3} a_jb_j = \sum_{i=1}^{3} a_ib_i = a_1b_1 + a_2b_2 + a_3b_3$, where the i, j and n are examples of the dummy indices that we sum over. The summation over repeated indices in vectorial and tensorial expressions is called the *Einstein summation convention* and simply saves us from having to write the summation sign over and over.

Another type of vector is the vector operator that we call the *gradient operator* that is defined

$$\nabla = \hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z} \mathrel{\widehat{=}} \text{gradient operator (a vector operator).} \tag{1.3}$$

For example, if $\psi(\boldsymbol{r}) = \psi(x, y, z)$ is some scalar field, then the gradient of ψ is

$$\nabla\psi = \hat{\boldsymbol{x}}\,\frac{\partial\psi}{\partial x} + \hat{\boldsymbol{y}}\,\frac{\partial\psi}{\partial y} + \hat{\boldsymbol{z}}\,\frac{\partial\psi}{\partial z}. \tag{1.4}$$

and is a vector that we can also write $\nabla\psi = \hat{\boldsymbol{x}}_i\partial\psi/\partial x_i$ using the summation convention. The gradient vector $\nabla\psi$ is oriented in the direction that the scalar field ψ is increasing the most rapidly and the amplitude $|\nabla\psi|$ gives the rate of that maximum increase.

For a vector field $\boldsymbol{a} = a_i\hat{\boldsymbol{x}}_i$, we call the a_i the *scalar components of the vector* and call the unit vectors $\hat{\boldsymbol{x}}_i$ in each direction i the *base vectors*. Note that a vector at some point in space is an arrow with a length and is completely independent of the coordinate system we use to describe it. So $\boldsymbol{a}$ happily exists as the same arrow and does not change if we rotate the coordinate system. Note, however, that the scalar components of the vector a_i will change as we rotate our coordinate system (alter the orientations of the base vectors) or switch to another coordinate system such as cylindrical coordinates.

Some authors put an arrow above a symbol to denote that it is a vector field, i.e., $\vec{a}$. When working in typed text, we always use a bold-face symbol to denote a vector, i.e., $\boldsymbol{a}$. When writing by hand, we have elected not to use an arrow over a symbol but instead use a squiggly underscore, i.e., $\underset{\sim}{a}$.

You are free to develop your own vectorial and tensorial notation when writing by hand but using squiggly underscores for vectors and tensors has served me well over a long career. Note that if you do not use some type of notation to denote that a symbol is a vector or tensor, you will be in a constant state of confusion when manipulating the fields of continuum physics.

Second-Order Tensor Fields: A field quantity that acts as the proportionality between two vector fields that are related to each other at each point in space is called a *second-order tensor* field (can equivalently be called a "second-rank" tensor). Another word that is synonymous to second-order tensor is *dyad* or *dyadic*. We write a second-order tensor field as

$$\begin{aligned}\boldsymbol{T}(\boldsymbol{r}) &\mathrel{\widehat{=}} \text{a second-order tensor field defined at each point } \boldsymbol{r}\\ &= T_{xx}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + T_{xy}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} + T_{xz}\hat{\boldsymbol{x}}\hat{\boldsymbol{z}}\\ &\quad + T_{yx}\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} + T_{yy}\hat{\boldsymbol{y}}\hat{\boldsymbol{y}} + T_{yz}\hat{\boldsymbol{y}}\hat{\boldsymbol{z}} \qquad (1.5)\\ &\quad + T_{zx}\hat{\boldsymbol{z}}\hat{\boldsymbol{x}} + T_{zy}\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} + T_{zz}\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\\ &= T_{ij}\,\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j \quad \text{(summation over repeated indices assumed).} \qquad (1.6)\end{aligned}$$

Just like the vector $\boldsymbol{a} = a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}$ is the sum of three vectors in the three coordinate directions, so the second-order tensor $\boldsymbol{T}$ is the sum of nine second-order tensors as made explicit in Eq. (1.5). The T_{ij} are the scalar components of the second-order tensor and the various base vector pairs $\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ for the various possible i and j are what we call second-order tensors. And just like we can write a vector in the array format $\boldsymbol{a} = (a_x, a_y, a_z)$, so can we write a second-order tensor as

$$\boldsymbol{T} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz}\\ T_{yx} & T_{yy} & T_{yz}\\ T_{zx} & T_{zy} & T_{zz}\end{pmatrix}. \qquad (1.7)$$

So a second-order tensor can be represented as a matrix. Much of what you learned about matrices in linear algebra applies to how we use second-order tensors. The main difference between a matrix and a second-order tensor is that although a matrix may have any dimension $(N \times M)$ and corresponds to any proportionality between an M and N dimensioned vector (first-order) array, a second-order tensor is a field quantity distributed through three-dimensional space and is always a (3×3) matrix in three-dimensional space and is a physical field that is always the proportionality between two vector fields that each have clear physical meaning as will be demonstrated repeatedly throughout this book.

An example of a second-order tensor is two vector fields that sit side by side to each other in an expression without a scalar or vector product (that are defined in an upcoming section) between them:

$$\boldsymbol{ab} = \left(a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right)\left(b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right) \qquad (1.8)$$

$$\begin{aligned}&= a_xb_x\,\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + a_xb_y\,\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} + a_xb_z\,\hat{\boldsymbol{x}}\hat{\boldsymbol{z}}\\ &\quad + a_yb_x\,\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} + a_yb_y\,\hat{\boldsymbol{y}}\hat{\boldsymbol{y}} + a_yb_z\,\hat{\boldsymbol{y}}\hat{\boldsymbol{z}} \qquad (1.9)\\ &\quad + a_zb_x\,\hat{\boldsymbol{z}}\hat{\boldsymbol{x}} + a_zb_y\,\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} + a_zb_z\,\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\\ &= a_ib_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j \quad \text{(summation over repeated indices as always).} \qquad (1.10)\end{aligned}$$

It is convenient to construct the 3×3 matrix representing $\boldsymbol{ab}$ as the matrix product between $\boldsymbol{a}$ written as a 3×1 array and $\boldsymbol{b}$ written as a 1×3 array, which corresponds to the multiplications of Eq. (1.8):

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\boldsymbol{ab} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \left(b_x, b_y, b_z\right) = \begin{pmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{pmatrix}. \tag{1.11}$$

When two vectors sit next to each other to form a second-order tensor, it is common to call that product the *tensor product* or *dyadic product*, even if we will not employ these words outside of this paragraph. Some authors in the engineering literature introduce a special symbol $\otimes$ to denote the tensor product, i.e., $\boldsymbol{a} \otimes \boldsymbol{b} \,\widehat{=}\, \boldsymbol{ab}$. So for the tensor product between the base vectors in any second-order or higher-order tensorial expression, these authors write, for example, $\hat{\boldsymbol{x}} \otimes \hat{\boldsymbol{y}}$ to represent what most authors write more simply as $\hat{\boldsymbol{x}}\hat{\boldsymbol{y}}$. The extra symbol $\otimes$ uses space on the page without providing any needed clarification, which is why we do not use it.

Another example of a second-order tensor is the gradient of a vector field. Working in Cartesian coordinates where derivatives of base vectors are zero, we have

$$\nabla \boldsymbol{a} = \left(\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right)\left(a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right) \tag{1.12}$$

$$= \frac{\partial a_x}{\partial x}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + \frac{\partial a_y}{\partial x}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} + \ldots = \frac{\partial a_j}{\partial x_i}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j, \tag{1.13}$$

which can again be written in array form as

$$\nabla \boldsymbol{a} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \left(a_x, a_y, a_z\right) = \begin{pmatrix} \frac{\partial a_x}{\partial x} & \frac{\partial a_y}{\partial x} & \frac{\partial a_z}{\partial x} \\ \frac{\partial a_x}{\partial y} & \frac{\partial a_y}{\partial y} & \frac{\partial a_z}{\partial y} \\ \frac{\partial a_x}{\partial z} & \frac{\partial a_y}{\partial z} & \frac{\partial a_z}{\partial z} \end{pmatrix}. \tag{1.14}$$

We emphasize that we get these simple expressions for the components of $\nabla \boldsymbol{a}$ only in Cartesian coordinates where derivatives of the base vectors are zero because the base vectors in Cartesians are spatially uniform. When the components of $\nabla \boldsymbol{a}$ are written out in curvilinear coordinates (cylindrical, spherical, etc.) in which the base vectors themselves vary with position in space and thus have nonzero spatial derivatives, the result of performing $\nabla \boldsymbol{a}$ is more complicated. The expressions for $\nabla \boldsymbol{a}$ in arbitrary orthogonal curvilinear coordinates, cylindrical coordinates, and spherical coordinates are all given in Section 1.8.6.

Just like with a matrix, we can talk about the transpose of a second-order tensor $\boldsymbol{T} = T_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ and write

$$\boldsymbol{T}^T \,\widehat{=}\, \text{the transpose of } \boldsymbol{T}$$

$$= T_{ij}\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i = T_{ji}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j. \tag{1.15}$$

Thus to perform the transpose, we can either flip the indices on the scalar components $T_{ij} \rightarrow T_{ji}$ of the tensor or flip the position of the two base vectors as they sit side by side.

Note that like with a vector, a tensor $\boldsymbol{T}$ exists at a point and is independent of the coordinate system. If we rotate or change coordinate systems, $\boldsymbol{T}$ does not change. However, the scalar components of the tensor T_{ij} will change as we change the coordinates because the base vectors $\hat{\boldsymbol{x}}_i$ are changing. When working in typed text, we always denote a second-order tensor with bold type. When we write a second-order tensor by hand, we use two squiggly underscores $\boldsymbol{T} = \underset{\approx}{T}$.

Higher-Order Tensor Fields: The generalization to higher-order tensors is straightforward. A third-order tensor is written

$$ {}_3\boldsymbol{P} = P_{ijk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k \tag{1.16}$$

a fourth-order tensor as

$$ {}_4\boldsymbol{Q} = Q_{ijkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.17}$$

and so on for still higher-order tensors. Summation over each index is again assumed.

If, for example, a second-order tensor $\boldsymbol{A}$ happens to sit next to two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ we would have the fourth-order tensor

$$ \boldsymbol{Aab} = A_{ij}a_k b_l\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l. \tag{1.18}$$

In general, we have $\boldsymbol{Aab} \neq \boldsymbol{Aba} \neq \boldsymbol{aAb} \neq \boldsymbol{bAa} \neq \boldsymbol{abA} \neq \boldsymbol{baA}$, so the order, from left to right, in which tensorial expressions sit next to each other to form higher-order tensors is very important.

The transpose of higher-order tensors must be specified by the way in which the base vectors are moved around relative to each other in the desired transpose operation. So, for example, for the fourth-order tensor ${}_4\boldsymbol{Q} = Q_{ijkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l$, we can define transpose operations such as

$$ {}_4\boldsymbol{Q}^{\overset{T}{2134}} = Q_{ijkl}\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l = Q_{jikl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.19}$$

$$ {}_4\boldsymbol{Q}^{\overset{T}{1243}} = Q_{ijkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_k = Q_{ijlk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.20}$$

$$ {}_4\boldsymbol{Q}^{\overset{T}{3412}} = Q_{ijkl}\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j = Q_{klij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l \tag{1.21}$$

and so on. There are $4! - 1 = 23$ such transposes for a fourth-order tensor, that is, there are $4! - 1$ different ways of placing the four base vectors next to each other that are different than in the nontransposed form. Similarly, an nth-order tensor would have $n! - 1$ different possible transpose operations; so a second-order tensor has only one way to write the transpose.

We write an nth-order tensor ${}_n\boldsymbol{Q}$ by hand as ${}_n\underset{\sim}{Q}$ for $n > 2$.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

1.3 Some Acts

In tensor calculus, just like in vector calculus, we define two commonly employed types of products between vectors and tensors called the *scalar product* and the *vector product*.

Scalar Products: A scalar product between two vector fields $\boldsymbol{a}$ and $\boldsymbol{b}$ that have an angle θ between them at each point as depicted in Fig. 1.2 is the product of the amplitude of the two vectors after one of the two vectors is projected into the direction of the other vector. The scalar product $\boldsymbol{a} \cdot \boldsymbol{b}$ between two vectors is a scalar and is denoted with a dot sitting between the vectors and is defined by the following rule

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta, \quad \text{where} \quad |\boldsymbol{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}. \tag{1.22}$$

So $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ if $\boldsymbol{a} \perp \boldsymbol{b}$, which means that $\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{y}} = 0$ and $\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{z}} = 0$, but $\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{x}} = 1$, etc. Using these rules, we thus have

$$\begin{aligned} \boldsymbol{a} \cdot \boldsymbol{b} &= \left[a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right] \cdot \left[b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right] \\ &= a_x b_x + a_y b_y + a_z b_z \\ &= a_i b_i. \end{aligned} \tag{1.23}$$

The scalar product is also called the *dot product* or the *inner product*.

What if vector field $\boldsymbol{a}$ is related to vector field $\boldsymbol{b}$ at some point in space? How do you obtain $\boldsymbol{a}$ given $\boldsymbol{b}$? That is what a second-order tensor such as $\boldsymbol{T} = T_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ does for us once we introduce the scalar product:

$$\begin{aligned} \boldsymbol{a} &= \boldsymbol{T} \cdot \boldsymbol{b} \\ &= \begin{pmatrix} & T_{xx}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}} & + & T_{xy}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}} & + & T_{xz}\hat{\boldsymbol{x}}\hat{\boldsymbol{z}} \\ + & T_{yx}\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} & + & T_{yy}\hat{\boldsymbol{y}}\hat{\boldsymbol{y}} & + & T_{yz}\hat{\boldsymbol{y}}\hat{\boldsymbol{z}} \\ + & T_{zx}\hat{\boldsymbol{z}}\hat{\boldsymbol{x}} & + & T_{zy}\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} & + & T_{zz}\hat{\boldsymbol{z}}\hat{\boldsymbol{z}} \end{pmatrix} \cdot \left(b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right) \end{aligned} \tag{1.24}$$

$$\begin{aligned} &= \left(T_{xx}\, b_x + T_{xy}\, b_y + T_{xz}\, b_z\right) \hat{\boldsymbol{x}} + \left(T_{yx}\, b_x + T_{yy}\, b_y + T_{yz}\, b_z\right) \hat{\boldsymbol{y}} \\ &\quad + \left(T_{zx}\, b_x + T_{zy}\, b_y + T_{zz}\, b_z\right) \hat{\boldsymbol{z}} \end{aligned} \tag{1.25}$$

$$= \left(T_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\right) \cdot \left(b_k\hat{\boldsymbol{x}}_k\right) = T_{ij}b_k\hat{\boldsymbol{x}}_i \left(\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_k\right) = T_{ij}b_j\hat{\boldsymbol{x}}_i, \tag{1.26}$$

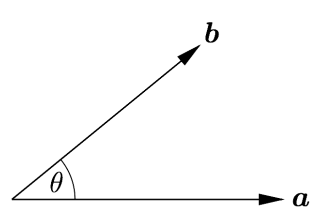

Figure 1.2 Two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ with an angle θ between them.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

where in the last line we used that $\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_k$ requires $k = j$. Using the familiar matrix multiplication for the scalar product, this can be written

$$\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}. \tag{1.27}$$

IMPORTANT: Second-order tensor fields are always maps between two vectors that are physically related to each other at a point $\boldsymbol{r}$. You cannot visualize directly a second-order tensor (or higher-order tensors) using your 3D sense of perception. But you can picture in your mind's eye the two vectors (arrows) that are related to each other at a point and thus imagine there is a mapping (second-order tensor) that takes the one vector to the other.

Note that throughout this entire book, we work exclusively in orthogonal coordinates where dot products are zero between the different base vectors of a coordinate system. It is possible, for example, in crystallography, to want to work in *skew* coordinate systems where the base vectors are not orthogonal to each other. Complicating ideas such as covariant and contravariant base vectors arise and the reader interested in tensor calculus in skew coordinates is directed toward specialized texts (e.g., Lebedev et al., 2010).

We can also speak of the *double-dot product* : between tensors, that in this book is defined

$$\boldsymbol{ab} : \boldsymbol{cd} = (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c}) \tag{1.28}$$

$$= \left(a_i \hat{\boldsymbol{x}}_i \cdot d_j \hat{\boldsymbol{x}}_j\right) \left(b_k \hat{\boldsymbol{x}}_k \cdot c_l \hat{\boldsymbol{x}}_l\right) \tag{1.29}$$

$$= a_i d_j b_k c_l \left(\hat{\boldsymbol{x}}_i \cdot \hat{\boldsymbol{x}}_j\right) \left(\hat{\boldsymbol{x}}_k \cdot \hat{\boldsymbol{x}}_l\right). \tag{1.30}$$

Other authors define the double-dot product as $(\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d})$. Either definition works if used consistently. We choose the convention of Eq. (1.28) so that when you see the : between vectors or base vectors, you perform the first dot product between the vectors that reside immediately on either side of the dot symbol and once that is done, perform the second dot product between the remaining vectors. This convention is the easiest to remember and is highly recommended. In writing out a tensorial expression such as given in Eq. (1.29), always use a different index for each base vector and associated coefficient. Because of the nature of the scalar product in orthogonal coordinate systems, we thus have $l = k$ and $j = i$ in Eq. (1.30) or

$$\boldsymbol{ab} : \boldsymbol{cd} = a_i b_k c_k d_i \quad \text{with summation over repeated indices} \tag{1.31}$$

$$= a_1 b_1 c_1 d_1 + a_2 b_1 c_1 d_2 + a_1 b_2 c_2 d_1 + a_2 b_2 c_2 d_2 \quad \text{in 2D.} \tag{1.32}$$

Note that for two second-order tensors $\boldsymbol{S}$ and $\boldsymbol{T}$, we have $\boldsymbol{S} \cdot \boldsymbol{T} = \left(\boldsymbol{T}^T \cdot \boldsymbol{S}^T\right)^T$ and that $\boldsymbol{S} \cdot \boldsymbol{T} \neq \boldsymbol{T} \cdot \boldsymbol{S}$ in general. For the double-dot product, however, we do have $\boldsymbol{S} : \boldsymbol{T} = \boldsymbol{T} : \boldsymbol{S}$ for any $\boldsymbol{S}$ and $\boldsymbol{T}$, where

$$\boldsymbol{S} : \boldsymbol{T} = S_{ij}\, \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j : T_{kl}\, \hat{\boldsymbol{x}}_k \hat{\boldsymbol{x}}_l \tag{1.33}$$

$$= S_{ij} T_{kl} \left(\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_k\right) \left(\hat{\boldsymbol{x}}_i \cdot \hat{\boldsymbol{x}}_l\right) \quad \text{which requires } l = i \text{ and } k = j \tag{1.34}$$

$$= S_{ij} T_{ji} \quad \text{with summation over repeated indices.} \tag{1.35}$$

Renaming the dummy indices gives $\boldsymbol{S} : \boldsymbol{T} = S_{ij}T_{ji} = S_{ji}T_{ij} = T_{ij}S_{ji} = \boldsymbol{T} : \boldsymbol{S}$.

The second-order *identity tensor* $\boldsymbol{I}$ is defined $\boldsymbol{I} = \delta_{ij}\, \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j$, where the δ_{ij} are called the *Kronecker coefficients* and are defined

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad \text{so that} \quad \boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.36}$$

Upon summing over the indices, we have $\boldsymbol{I} = \hat{\boldsymbol{x}}_1 \hat{\boldsymbol{x}}_1 + \hat{\boldsymbol{x}}_2 \hat{\boldsymbol{x}}_2 + \hat{\boldsymbol{x}}_3 \hat{\boldsymbol{x}}_3 = \hat{\boldsymbol{x}}\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}}\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}\hat{\boldsymbol{z}}$. The identity tensor works as follows: $\boldsymbol{A} \cdot \boldsymbol{I} = \boldsymbol{I} \cdot \boldsymbol{A} = \boldsymbol{A}$ for any second-order tensor $\boldsymbol{A}$. We further have that if the position vector is written $\boldsymbol{r} = x_j \hat{\boldsymbol{x}}_j$ in Cartesian coordinates, then $\boldsymbol{I} = \nabla \boldsymbol{r} = (\partial x_j / \partial x_i) \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j = \delta_{ij} \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j$.

A double-dot product with the identity tensor results in $\boldsymbol{A} : \boldsymbol{I} = A_{ij}\, \delta_{ji} = A_{ii} = \text{tr}\, \{\boldsymbol{A}\} = A_{11} + A_{22} + A_{33}$, which is called the *trace* of second-order tensor $\boldsymbol{A}$. The trace is the sum of the second-order tensor components along the diagonal, for example, $\boldsymbol{I} : \boldsymbol{I} = 3$ (in 3D). The double-dot product between two second-order tensors is the trace of the scalar (matrix) product of the two tensors, that is, $\boldsymbol{A} : \boldsymbol{B} = A_{ij}B_{ji} = \text{tr}\, \{\boldsymbol{A} \cdot \boldsymbol{B}\} = \text{tr}\, \{\boldsymbol{B} \cdot \boldsymbol{A}\}$.

We can extend the number of dot products we take between two higher-order tensors to as many as desired. So the *triple-dot product* $\overset{3}{:}$ between, say, two third-order tensors ${}_3\boldsymbol{S}$ and ${}_3\boldsymbol{T}$ can be defined

$$\begin{aligned} {}_3\boldsymbol{S} \overset{3}{:} {}_3\boldsymbol{T} &= S_{ijk} \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k \overset{3}{:} T_{lmn} \hat{\boldsymbol{x}}_l \hat{\boldsymbol{x}}_m \hat{\boldsymbol{x}}_n \\ &= S_{ijk} T_{lmn} (\hat{\boldsymbol{x}}_k \cdot \hat{\boldsymbol{x}}_l)(\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_m)(\hat{\boldsymbol{x}}_i \cdot \hat{\boldsymbol{x}}_n), \end{aligned}$$

which tells us that $n = i$, $m = j$ and $l = k$ so that the triple-dot product between two third-order tensors comes out to be

$${}_3\boldsymbol{S} \overset{3}{:} {}_3\boldsymbol{T} = S_{ijk} T_{kji} = \text{tr}\, \{{}_3\boldsymbol{S} : {}_3\boldsymbol{T}\} \tag{1.37}$$

with summation over repeated indices. We can extend such notation and definition to still higher-order dot products between still higher-order tensors.

Note that each dot product removes two base vectors from a tensorial expression. So without writing anything out, we know that a tensorial expression like ${}_8\boldsymbol{A} \overset{5}{:} {}_6\boldsymbol{B}$ is a fourth-order tensor, that is, the eighth-order tensor ${}_8\boldsymbol{A}$ contributes 8 base vectors to this expression and the sixth-order tensor ${}_6\boldsymbol{B}$ contributes 6 more base vectors but the 5 dot products remove 10 of those base vectors so that the result is a fourth-order tensor. As practice, we can write this lengthy example out to give

$$\begin{aligned} &{}_8\boldsymbol{A} \overset{5}{:} {}_6\boldsymbol{B} = \\ &A_{ijklmnop} B_{qrstuv} \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k (\hat{\boldsymbol{x}}_p \cdot \hat{\boldsymbol{x}}_q)(\hat{\boldsymbol{x}}_o \cdot \hat{\boldsymbol{x}}_r)(\hat{\boldsymbol{x}}_n \cdot \hat{\boldsymbol{x}}_s)(\hat{\boldsymbol{x}}_m \cdot \hat{\boldsymbol{x}}_t)(\hat{\boldsymbol{x}}_l \cdot \hat{\boldsymbol{x}}_u) \hat{\boldsymbol{x}}_v, \end{aligned} \tag{1.38}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

which tells us that $q = p$, $r = o$, $s = n$, $t = m$, and $u = l$ so that we obtain the fourth-order tensor

$$ {}_8\boldsymbol{A} \overset{5}{\vdots} {}_6\boldsymbol{B} = A_{ijklmnop} B_{ponmlv} \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k \hat{\boldsymbol{x}}_v \tag{1.39} $$

with summation implied over all the dummy (i.e., repeated) indices. In our development of continuum physics, we will not need to work with tensors higher than the sixth order or with more than three dot products between two tensors.

Vector Products: A vector product between two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is a vector that is perpendicular to the two vectors as depicted in Fig. 1.3 and that has an amplitude equal to the area of the parallelogram formed with the two vectors as sides. We have

$$ \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}; \quad \boldsymbol{c} \perp \text{ to both } \boldsymbol{a} \text{ and } \boldsymbol{b} \tag{1.40} $$

$$ |\boldsymbol{c}| = |\boldsymbol{a}||\boldsymbol{b}| \sin\theta \tag{1.41} $$

$$ \boldsymbol{a} \times \boldsymbol{b} = 0 \quad \text{if } \boldsymbol{a} \,||\, \boldsymbol{b}. \tag{1.42} $$

Use the right-hand rule to determine the sense of $\boldsymbol{a} \times \boldsymbol{b}$. Note that $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$. We can thus obtain

$$ \hat{\boldsymbol{x}} \times \hat{\boldsymbol{y}} = \hat{\boldsymbol{z}}; \quad \hat{\boldsymbol{x}} \times \hat{\boldsymbol{z}} = -\hat{\boldsymbol{y}}; \quad \hat{\boldsymbol{x}} \times \hat{\boldsymbol{x}} = 0 \tag{1.43} $$

and so forth for all the vector products between all base vectors. Using these rules, we can write

$$ \begin{aligned} \boldsymbol{a} \times \boldsymbol{b} &= \left(a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right) \times \left(b_x\hat{\boldsymbol{x}} + b_y\hat{\boldsymbol{y}} + b_z\hat{\boldsymbol{z}}\right) \\ &= a_x b_y \hat{\boldsymbol{z}} - a_x b_z \hat{\boldsymbol{y}} - a_y b_x \hat{\boldsymbol{z}} + a_y b_z \hat{\boldsymbol{x}} + a_z b_x \hat{\boldsymbol{y}} - a_z b_y \hat{\boldsymbol{x}} \\ &= \left(a_y b_z - a_z b_y\right) \hat{\boldsymbol{x}} - \left(a_x b_z - a_z b_x\right) \hat{\boldsymbol{y}} + \left(a_x b_y - a_y b_x\right) \hat{\boldsymbol{z}}. \end{aligned} \tag{1.44} $$

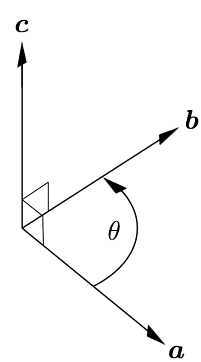

Figure 1.3 The vector product $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$ is a vector perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$ as determined by the right-hand rule.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

We can thus write the vector product using the matrix determinant in the following way:

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}, \text{ where } |\cdots| \text{ denotes } \textit{taking the determinant}. \tag{1.45}$$

Note that the vector product is also called the *cross product*.

Doing Vector (Cross) Products with Scalar (Dot) Products: It is convenient for proving identities involving the cross product to write a cross product in a way that only involves dot products. To do so, we introduce the *alternating* or *permutation* or *antisymmetric* or *Levi–Civita* (these are all synonyms) third-order tensor ${}_3\boldsymbol{\epsilon}$ that is defined

$${}_3\boldsymbol{\epsilon} = \epsilon_{ijk}\, \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k \tag{1.46}$$

with scalar components called the *Levi–Civita coefficients* given by

$$\epsilon_{ijk} = \begin{cases} +1 & \text{for counterclockwise index positions: 123, 231, 312} \\ -1 & \text{for clockwise positions: 132, 321, 213} \\ \;\;0 & \text{for every other index combination,} \end{cases} \tag{1.47}$$

where this cyclic ordering of the indices can be remembered using the mnemonic device of Fig. 1.4. To perform the cross product with the Levi–Civita tensor, we do the following

$$\begin{aligned} \boldsymbol{a} \times \boldsymbol{b} &= {}_3\boldsymbol{\epsilon} : \boldsymbol{ba} = -{}_3\boldsymbol{\epsilon} : \boldsymbol{ab} = -\epsilon_{ijk}\, \hat{\boldsymbol{x}}_i\, \hat{\boldsymbol{x}}_j\, \hat{\boldsymbol{x}}_k : a_l\, \hat{\boldsymbol{x}}_l\, b_m\, \hat{\boldsymbol{x}}_m \\ &= -\epsilon_{ijk}\, a_l b_m\, \hat{\boldsymbol{x}}_i \left(\hat{\boldsymbol{x}}_k \cdot \hat{\boldsymbol{x}}_l\right) \left(\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_m\right) \\ &= -\epsilon_{ijk}\, a_k b_j\, \hat{\boldsymbol{x}}_i & (1.48) \\ &= \hat{\boldsymbol{x}}_1 \left[a_2 b_3 - a_3 b_2\right] + \hat{\boldsymbol{x}}_2 \left[a_3 b_1 - a_1 b_3\right] + \hat{\boldsymbol{x}}_3 \left[a_1 b_2 - a_2 b_1\right] & (1.49) \end{aligned}$$

which is identical to Eq. (1.44). Note that the double-dot product with the Levi–Civita tensor uses our double-dot convention of Eq. (1.28), which explains the minus sign in Eq. (1.48) in comparison to other authors who use the alternative but less-intuitive definition of the double-dot product.

If the Levi–Civita tensor is double dotted into the second-order tensor $\boldsymbol{A}$, we obtain the vector

$${}_3\boldsymbol{\epsilon} : \boldsymbol{A} = \hat{\boldsymbol{x}}_1 \left(A_{32} - A_{23}\right) + \hat{\boldsymbol{x}}_2 \left(A_{13} - A_{31}\right) + \hat{\boldsymbol{x}}_3 \left(A_{21} - A_{12}\right). \tag{1.50}$$

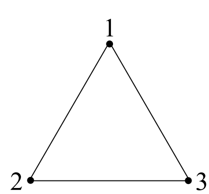

Figure 1.4 Mnemonic triangle showing the counterclockwise ordering of the three Levi–Civita indices

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

So if the tensor $\boldsymbol{A}$ is symmetric, then ${}_3\boldsymbol{\epsilon} : \boldsymbol{A} = 0$. Any second-order tensor can be separated into symmetric and antisymmetric portions $\boldsymbol{A} = \boldsymbol{A}^{(s)} + \boldsymbol{A}^{(a)}$, where

$$\boldsymbol{A}^{(s)} = \frac{1}{2}\left(\boldsymbol{A} + \boldsymbol{A}^T\right), \tag{1.51}$$

$$\boldsymbol{A}^{(a)} = \frac{1}{2}\left(\boldsymbol{A} - \boldsymbol{A}^T\right). \tag{1.52}$$

So the antisymmetric portion of any second-order tensor has zeroes along the diagonal and off-diagonal components that are "antisymmetric" $A^{(a)}_{ij} = -A^{(a)}_{ji}$. So the operation ${}_3\boldsymbol{\epsilon} : \boldsymbol{A} = {}_3\boldsymbol{\epsilon} : \boldsymbol{A}^{(a)}$ given by Eq. (1.50) involves only the antisymmetric portion of $\boldsymbol{A}$. This is because the Levi–Civita coefficients are anti-symmetric, that is, $\epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{jik}$. You can prove as an end-of-chapter exercise that the double-dot product of any tensor that is antisymmetric in the last two base vectors with any tensor that is symmetric in the first two base vectors gives zero.

The antisymmetric nature of ${}_3\boldsymbol{\epsilon}$ can further be seen by dotting it into any vector $\boldsymbol{a}$ to obtain the second-order tensor

$${}_3\boldsymbol{\epsilon} \cdot \boldsymbol{a} = \epsilon_{ijk} a_k \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}, \tag{1.53}$$

which is the definition of an antisymmetric second-order tensor. So the dot product of the third-order Levi–Civita tensor with any vector always produces an antisymmetric second-order tensor.

There is a useful identity involving the Levi–Civita coefficients

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}, \tag{1.54}$$

where the δ_{ij} are the Kronecker coefficients. For the left-hand side to be nonzero, we need $j, k, l, m \neq i$ as well as both $j \neq k$ and $l \neq m$. If we take $j = l$ and $k = m$, then $\epsilon_{ijk}\epsilon_{ilm} = \epsilon_{ijk}\epsilon_{ijk} = 1$. If we take $j = m$ and $k = l$, then $\epsilon_{ijk}\epsilon_{ilm} = \epsilon_{ijk}\epsilon_{jik} = -1$. This suite of conditions is exactly satisfied by the right-hand side of Eq. (1.54). The identity of Eq. (1.54) allows us to prove relations that involve two cross products.

As an example, express the double-cross product between three vectors as

$$\boldsymbol{a} \times \boldsymbol{b} \times \boldsymbol{c} = -\boldsymbol{a} \times ({}_3\boldsymbol{\epsilon} : \boldsymbol{bc}) = {}_3\boldsymbol{\epsilon} : \boldsymbol{a}\,({}_3\boldsymbol{\epsilon} : \boldsymbol{bc}), \tag{1.55}$$

$$= \epsilon_{ijk} a_k \epsilon_{jlm} b_m c_l \hat{\boldsymbol{x}}_i. \tag{1.56}$$

Exchanging the dummy indices i and j and noting that $\epsilon_{jik} = -\epsilon_{ijk}$ gives

$$\boldsymbol{a} \times \boldsymbol{b} \times \boldsymbol{c} = -\epsilon_{ijk}\epsilon_{ilm} a_k b_m c_l \hat{\boldsymbol{x}}_j. \tag{1.57}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

The identity of Eq. (1.54) then results in

$$\boldsymbol{a} \times \boldsymbol{b} \times \boldsymbol{c} = \left(\delta_{jm}\delta_{kl} - \delta_{jl}\delta_{km}\right) a_k b_m c_l \hat{\boldsymbol{x}}_j, \tag{1.58}$$

$$= \left(a_k b_j c_k - a_k b_k c_j\right) \hat{\boldsymbol{x}}_j, \tag{1.59}$$

$$= \boxed{(\boldsymbol{a} \cdot \boldsymbol{c})\, \boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{c}.} \tag{1.60}$$

The Levi–Civita third-order tensor allows us to rewrite cross products in terms of dot products, which simplifies obtaining vectorial and tensorial identities that involve the cross product.

1.4 The Integral Theorems

There are a variety of extremely useful theorems that involve the combined operations of differentiation and integration (the derivative and antiderivative). We will use these theorems over and over again in our development and manipulation of the rules of continuum physics.

Fundamental Theorem of 3D Calculus: For some volumetric region Ω bounded by the closed surface $\partial\Omega$ that has an outward normal $\boldsymbol{n}$ at each point of $\partial\Omega$ as depicted in Fig. 1.5, we have

$$\boxed{\int_{\Omega} \nabla\psi(\boldsymbol{r})\, \mathrm{d}^3\boldsymbol{r} = \int_{\partial\Omega} \boldsymbol{n}\psi(\boldsymbol{r})\, \mathrm{d}^2\boldsymbol{r},} \quad \textit{fundamental theorem of 3D calculus} \tag{1.61}$$

where ψ is a scalar, vector, or tensor field of any order. If ψ is either a vector or tensor, it is necessary that it appears in the position given within the integrand on the right-hand side of Eq. (1.61). Although this theorem is used repeatedly in physics, for some peculiar reason it is rarely presented in vector calculus texts intended for the undergraduate level. As a guided exercise at the end of the chapter, you can prove Eq. (1.61) rather easily. Equation (1.61) is the 3D generalization of the *fundamental theorem of 1D calculus*

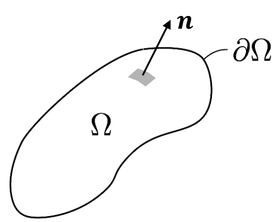

Figure 1.5 An arbitrary region Ω bounded by the closed surface $\partial\Omega$.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\boxed{\int_a^b \frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\,\mathrm{d}x = \psi(b) - \psi(a).} \qquad \textit{fundamental theorem of 1D calculus} \tag{1.62}$$

In the above volume and surface integrals, we have employed the notation that $\mathrm{d}^3\boldsymbol{r} = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \mathrm{d}V$ denotes a volume element and $\mathrm{d}^2\boldsymbol{r} = \mathrm{d}S$ denotes a surface element. With $\psi \to \boldsymbol{a}$ in the fundamental theorem of 3D calculus, where $\boldsymbol{a}$ is a vector or tensor field of any order, taking the trace over the two base vectors gives

$$\int_\Omega \nabla \cdot \boldsymbol{a}\,\mathrm{d}^3\boldsymbol{r} = \int_{\partial\Omega} \boldsymbol{n} \cdot \boldsymbol{a}\,\mathrm{d}^2\boldsymbol{r} \qquad \textit{divergence theorem} \tag{1.63}$$

that is also called *Gauss' theorem*.

One may similarly obtain

$$\int_\Omega \nabla \times \boldsymbol{a}\,\mathrm{d}^3\boldsymbol{r} = \int_{\partial\Omega} \boldsymbol{n} \times \boldsymbol{a}\,\mathrm{d}^2\boldsymbol{r} \qquad \textit{curl theorem} \tag{1.64}$$

which is not Stokes' theorem even if it involves the *curl* operation $\nabla \times \boldsymbol{a}$ defined in Section 1.6.

Stokes' Theorem: For some finite open (possibly curved) surface S having normal $\boldsymbol{n}$ at each point on S and bounded by a closed contour Γ as depicted in Fig. 1.6, we have

$$\boxed{\int_S \boldsymbol{n} \cdot (\nabla \times \boldsymbol{a})\,\mathrm{d}S = \oint_\Gamma \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{l}} \qquad \textit{Stokes' theorem} \tag{1.65}$$

where $\mathrm{d}\boldsymbol{l}$ is the infinitesimal length vector tangent to points on Γ and $\oint_\Gamma$ means that we start the integral on the closed contour Γ at one point, go around the contour in the counterclockwise direction and finish the integral at that same point.

When applied to a plane, Stokes' theorem is equivalent to a theorem that is usually called *Green's theorem*. So taking the open surface S to reside in the x, y plane and bounded by the closed curve Γ and considering two functions of (x, y) that we call $P(x, y)$ and $Q(x, y)$, Green's theorem states

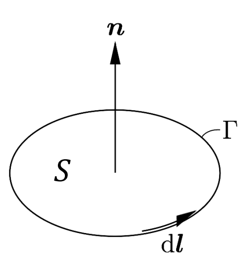

Figure 1.6 An open surface S with a normal vector $\boldsymbol{n}$ at each point and bounded by the closed contour Γ.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\int_S \left(\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y} \right) \mathrm{d}x\mathrm{d}y = \oint_\Gamma \left[P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y \right]. \tag{1.66}$$

By substituting $P(x, y) = a_x(x, y)$ and $Q(x, y) = a_y(x, y)$, Green's theorem becomes Stokes' theorem. You can prove Green's theorem as an easy guided, end-of-chapter exercise. Green's theorem will also be used in Chapter 11 to prove Cauchy's theorem, which is the foundation for all contour-integration methods.

Differentiating under the Integral Sign: Another class of integral theorems involves time differentiation of spatial integrals when the limits of the integral domain are themselves variable in time.

In 3D, let's imagine a spatial integral domain $\Omega(t)$ whose enclosing boundary $\partial\Omega(t)$ is changing through time t because each point of the boundary is moving with a velocity $\boldsymbol{v}(\boldsymbol{r}, t)$. In this scenario, the time derivative of a volume integral over the time-variable domain $\Omega(t)$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} \psi(\boldsymbol{r}, t)\,\mathrm{d}^3\boldsymbol{r} = \int_{\Omega(t)} \frac{\partial \psi(\boldsymbol{r}, t)}{\partial t}\,\mathrm{d}^3\boldsymbol{r} + \int_{\partial\Omega(t)} \boldsymbol{n} \cdot \boldsymbol{v}(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t)\,\mathrm{d}^2\boldsymbol{r}, \tag{1.67}$$

which is called the *Reynolds transport theorem*. Here, ψ can be a scalar, vector, or tensor field of any order. Applying the divergence theorem to the surface integral gives the alternative expression

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega(t)} \psi(\boldsymbol{r}, t)\,\mathrm{d}^3\boldsymbol{r} = \int_{\Omega(t)} \left\{ \frac{\partial \psi(\boldsymbol{r}, t)}{\partial t} + \nabla \cdot [\boldsymbol{v}(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t)] \right\} \mathrm{d}^3\boldsymbol{r}. \tag{1.68}$$

If the field ψ is a vector or tensor, it is very important that it is placed after the velocity field in the tensorial expression that the divergence operates on. If the boundary $\partial\Omega$ is not moving with a velocity $\boldsymbol{v}$, then the time derivative passes through the volume integral and acts directly on the integrand ψ and there is no surface integral term or divergence term. If $\psi = 1$, then $V(t) = \int_{\Omega(t)} \mathrm{d}^3\boldsymbol{r}$ is the evolving volume of the domain $\Omega(t)$ and the Reynolds transport theorem gives $\mathrm{d}V(t)/\mathrm{d}t = \int_{\partial\Omega(t)} \boldsymbol{n} \cdot \boldsymbol{v}\,\mathrm{d}^2\boldsymbol{r}$, which is a self-evident fact and a result we will use in Chapter 4.

The Reynolds transport theorem of Eq. (1.67) is the 3D generalization of the rule for time differentiating 1D spatial integrals over a time-variable domain

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{a(t)}^{b(t)} \psi(x, t)\,\mathrm{d}x = \int_{a(t)}^{b(t)} \frac{\partial \psi(x, t)}{\partial t}\,\mathrm{d}x + \frac{\mathrm{d}b(t)}{\mathrm{d}t}\,\psi(b(t), t) - \frac{\mathrm{d}a(t)}{\mathrm{d}t}\,\psi(a(t), t), \tag{1.69}$$

which is called the *Leibniz rule*. Note that $\mathrm{d}b(t)/\mathrm{d}t$ is the velocity at which the domain limit $x = b$ is moving in the $+x$ direction and similarly for $\mathrm{d}a(t)/\mathrm{d}t$.

1.5 Divergence of Vector (and Tensor) Fields

The divergence operation is the dot product between the gradient operator and either a vector or tensor

$$\nabla \cdot \boldsymbol{a} = \left[\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right] \cdot \left[a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}\right]. \tag{1.70}$$

In carrying out the products and derivatives, note that in Cartesian coordinates the base vectors are uniform constants and have the property that $\partial\hat{\boldsymbol{x}}_j/\partial x_i = 0$. But this is not true for other base vectors in curvilinear coordinates (e.g., cylindrical and spherical). Since the derivatives of the base vectors are zero in Cartesian coordinates, we have

$$\nabla \cdot \boldsymbol{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} = \frac{\partial a_i}{\partial x_i} \tag{1.71}$$

with summation over the index i as always. In Section 1.8.6, we will investigate how this and other tensor-calculus operations involving the ∇ operator are different in orthogonal curvilinear coordinates. Our goal in the present section is specifically to provide physical understanding about what the divergence of vector (and tensor) fields is telling us about the field.

To do so, construct a region Ω around the point $\boldsymbol{r}$ that has a tiny volume $\delta V = \int_\Omega \mathrm{d}V$, where the δ means "tiny." The region is surrounded by the closed surface $\partial\Omega$. The divergence theorem says that:

$$\frac{1}{\delta V}\int_\Omega \nabla \cdot \boldsymbol{a}\, \mathrm{d}V = \frac{1}{\delta V}\int_{\partial\Omega} \boldsymbol{n} \cdot \boldsymbol{a}\, \mathrm{d}S. \tag{1.72}$$

In the limit as $\delta V \to 0$, $\nabla \cdot \boldsymbol{a} \approx$ constant in Ω so

$$\begin{aligned}\nabla \cdot \boldsymbol{a} &= \lim_{\delta V\to 0} \frac{1}{\delta V}\int_\Omega \nabla \cdot \boldsymbol{a}\, \mathrm{d}V \\ &= \lim_{\delta V\to 0} \frac{1}{\delta V}\int_{\partial\Omega} \boldsymbol{n} \cdot \boldsymbol{a}\, \mathrm{d}S,\end{aligned} \tag{1.73}$$

which represents the accumulation of the physical quantity carried by $\boldsymbol{a}$. Imagine the vector $\boldsymbol{a}$ as a flux of some type carrying a physical quantity with it. If the flux into a small volume element is different than the flux out of the element, which is what the surface integral of Eq. (1.73) quantifies, then the divergence of this flux is nonzero. The divergence of a vector field is thus associated with the idea of accumulation (if the divergence is negative) or depletion (if the divergence is positive) in a tiny region surrounding the point in question.

To conclude: *If $|\boldsymbol{n} \cdot \boldsymbol{a}|$ is larger on one side of the small element Ω compared to the other, the divergence is nonzero. If the divergence is positive, the physical quantity carried by $\boldsymbol{a}$ is depleting in Ω, while if the divergence is negative, the physical quantity is accumulating.* This is shown in Fig. 1.7.

There are two ways we commonly use to visualize a vector field as depicted in Fig. 1.8. In the first, we place a vector at each point in space. In the second, called *current lines* (or flow lines), the direction of the field at a point is tangent to the current line at that point

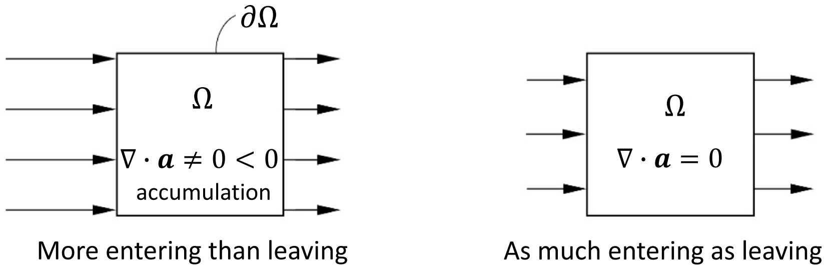

Figure 1.7 Figure showing that $\nabla \cdot \boldsymbol{a} < 0$ quantifies the accumulation of the physical quantity carried by the "flux" $\boldsymbol{a}$ due to more of the quantity fluxing into a volume than fluxing out. If the flux arrows leaving the volume are greater than those entering, the quantity is depleting and $\nabla \cdot \boldsymbol{a} > 0$. If the flux in equals the flux out, then there is no accumulation or depletion and $\nabla \cdot \boldsymbol{a} = 0$.

and the amplitude is given by the density of current lines in the neighborhood surrounding the point.

In the current line approach, the places where $\nabla \cdot \boldsymbol{a} \neq 0$ are always associated with the start of a new line as shown in Fig. 1.9. In particular, if we imagine the electric field $\boldsymbol{E}(\boldsymbol{r})$ around a point charge q, all field lines start at the charge location where $\nabla \cdot \boldsymbol{E} \neq 0$ but at all other points where there is no point charge, field lines are not being created and $\nabla \cdot \boldsymbol{E} = 0$.

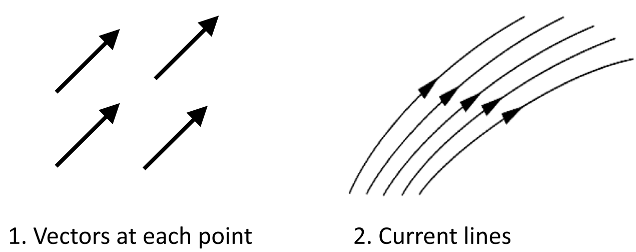

Figure 1.8 Two approaches that are commonly employed for picturing a vector field.

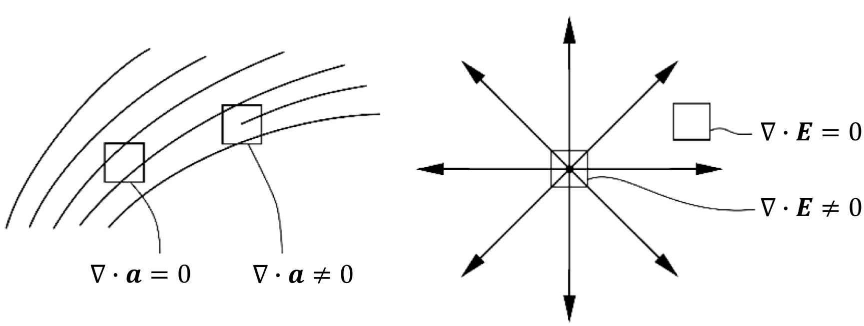

Figure 1.9 Places where $\nabla \cdot \boldsymbol{a} \neq 0$ are associated with the start of a current line.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

What about the divergence of a second-order tensor $\boldsymbol{T}$? Working in Cartesian coordinates where derivatives of the base vectors are zero (the curvilinear-coordinate expression is given later), we obtain $\nabla \cdot \boldsymbol{T}$ as

$$\nabla \cdot \boldsymbol{T} = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right] \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix}$$

$$= \Bigg(\underbrace{\frac{\partial T_{xx}}{\partial x} + \frac{\partial T_{yx}}{\partial y} + \frac{\partial T_{zx}}{\partial z}}_{x \text{ component}}, \underbrace{\frac{\partial T_{xy}}{\partial x} + \frac{\partial T_{yy}}{\partial y} + \frac{\partial T_{zy}}{\partial z}}_{y \text{ component}}, \underbrace{\frac{\partial T_{xz}}{\partial x} + \frac{\partial T_{yz}}{\partial y} + \frac{\partial T_{zz}}{\partial z}}_{z \text{ component}} \Bigg) \tag{1.74}$$

$$= \left(\hat{\boldsymbol{x}}_i \frac{\partial}{\partial x_i}\right) \cdot \left(T_{jk}\, \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k\right) = \frac{\partial T_{jk}}{\partial x_i} \left(\hat{\boldsymbol{x}}_i \cdot \hat{\boldsymbol{x}}_j\right) \hat{\boldsymbol{x}}_k = \frac{\partial T_{ik}}{\partial x_i} \hat{\boldsymbol{x}}_k \mathrel{\widehat{=}} \nabla \cdot \boldsymbol{T}. \tag{1.75}$$

To visualize or intuit the meaning of the divergence of a second-order tensor, we can use our same device of considering a tiny element Ω of volume δV surrounding the point where the vector $\nabla \cdot \boldsymbol{T}$ is being evaluated and obtain (because δV is so small that $\nabla \cdot \boldsymbol{T}$ is uniform inside of Ω)

$$\nabla \cdot \boldsymbol{T} = \lim_{\delta V \to 0} \frac{1}{\delta V} \int_\Omega \nabla \cdot \boldsymbol{T}\, \mathrm{d}V = \lim_{\delta V \to 0} \frac{1}{\delta V} \int_{\partial\Omega} \boldsymbol{n} \cdot \boldsymbol{T}\, \mathrm{d}S. \tag{1.76}$$

Regardless of what the second-order tensor $\boldsymbol{T}$ actually represents, think of the vector $\boldsymbol{n} \cdot \boldsymbol{T}\, \mathrm{d}S$ as being a force acting on all points on $\partial\Omega$ so that if the force acting on one side of the element is larger than the force acting on the other side (which is what the surface integral allows for), there is a net force acting on the element characterized by $\nabla \cdot \boldsymbol{T} \neq 0$.

1.6 Curl of Vector (and Tensor) Fields

The curl is the vector product between the gradient operator and a vector (or tensor). In Cartesian coordinates, in which the derivatives of the base vectors are zero, we can calculate the curl of a vector field using the determinant rule

$$\nabla \times \boldsymbol{a} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix}$$

$$= \hat{\boldsymbol{x}} \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z}\right) - \hat{\boldsymbol{y}} \left(\frac{\partial a_z}{\partial x} - \frac{\partial a_x}{\partial z}\right) + \hat{\boldsymbol{z}} \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}\right). \tag{1.77}$$

The curl is also sometimes called the *rotation* for reasons explained next.

For the physical meaning of the curl operation, we use Stokes' theorem in the limit that the open surface S at a point, which has a normal $\boldsymbol{n}$, has a tiny area δA sufficiently small that $\nabla \times \boldsymbol{a}$ can be taken as a constant over S

$$\boldsymbol{n} \cdot \nabla \times \boldsymbol{a} = \lim_{\delta A \to 0} \frac{1}{\delta A} \int_S \boldsymbol{n} \cdot (\nabla \times \boldsymbol{a})\, \mathrm{d}S = \lim_{\delta A \to 0} \frac{1}{\delta A} \oint_\Gamma \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{l}. \tag{1.78}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

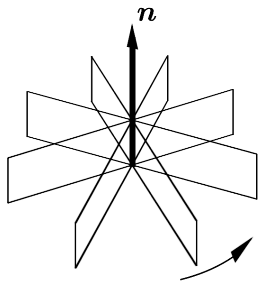

Figure 1.10 A waterwheel with flat blades and an axle in the $\boldsymbol{n}$ direction.

Thus, if the vector field $\boldsymbol{a}$ is tangent in places to the curve Γ and if this tangential component is larger on one side of S compared to the other so that integral over the closed line Γ is nonzero, then we will have $\nabla \times \boldsymbol{a} \neq 0$. We can associate this with the idea that the field has a rotation on S.

To visualize the curl, we use a *water wheel* consisting of an axle oriented in the direction $\boldsymbol{n}$ and having flat blades coming off the axle perpendicularly as depicted in Fig. 1.10. We immerse this wheel in our vector field $\boldsymbol{a}$ that we imagine to be the flow of water regardless of what $\boldsymbol{a}$ really corresponds to. We change the orientation of the axle of the wheel and observe how fast the wheel is moving if at all. The direction $\boldsymbol{n}$ of the axle at which the wheel turns the fastest in the counterclockwise direction gives the direction of $\nabla \times \boldsymbol{a}$, while the rate of rotation gives $|\nabla \times \boldsymbol{a}|$. The right-hand rule corresponds to the wheel rotating in the counterclockwise direction.

So to mentally investigate the curl (or rotation) associated with some field, we imagine the vector field to be a flow field and probe the field with our water wheel to see in what orientation the wheel turns the fastest if, indeed, it can turn at all, cf., Fig. 1.11. Note that

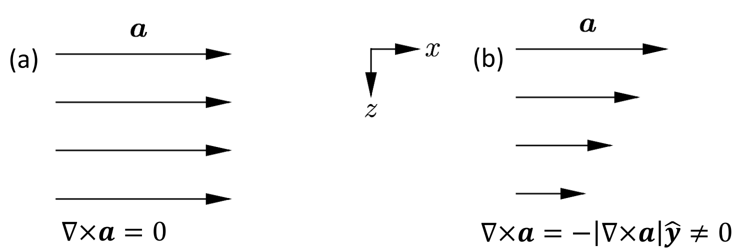

Figure 1.11 Use a water wheel to probe a field and see whether the wheel can turn: (a) a uniform vector field or a field with only longitudinal variation in the direction of the field cannot make a water wheel turn so that $\nabla \times \boldsymbol{a} = 0$; (b) a vector field that has spatial variation in a direction transverse to the field direction will always make a water wheel turn resulting in nonzero $\nabla \times \boldsymbol{a}$.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

a water wheel is not the same as a propellar. A propellar would turn even in case (a) of Fig. 1.11 corresponding to a uniform field.

> CONCLUSIONS: To conclude these discussions of how to understand and visualize nonzero values of the divergence and curl of a vector field, we can define *longitudinal* and *transverse* spatial variations of a vector field. Longitudinal variations are those in the same direction of the vector field and transverse variations are those in a direction perpendicular to the vector field. Longitudinal variations (the directional derivative $\hat{\boldsymbol{a}} \cdot \nabla \boldsymbol{a}$ in the direction of the vector field) lead to a nonzero divergence (unless the longitudinal variations in each direction sum to zero) but do not contribute to the curl. Similarly, transverse variations always lead to a nonzero curl but do not contribute to the divergence. The decomposition of a vector field into longitudinal and transverse variations will correspond to P-waves and S-waves, respectively, when we discuss elastic-wave propagation.

We can also perform the *curl* differential operation using the Levi–Civita tensor in the following way:

$$\nabla \times \boldsymbol{a} = -{}_3\boldsymbol{\epsilon} : \nabla \boldsymbol{a} = -\epsilon_{ijk}\, \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k : \hat{\boldsymbol{x}}_l \frac{\partial}{\partial x_l} a_m \hat{\boldsymbol{x}}_m = \boxed{-\epsilon_{ijk} \frac{\partial a_j}{\partial x_k} \hat{\boldsymbol{x}}_i} \tag{1.79}$$

$$= \hat{\boldsymbol{x}}_1 \left(\frac{\partial a_3}{\partial x_2} - \frac{\partial a_2}{\partial x_3} \right) + \hat{\boldsymbol{x}}_2 \left(\frac{\partial a_1}{\partial x_3} - \frac{\partial a_3}{\partial x_1} \right) + \hat{\boldsymbol{x}}_3 \left(\frac{\partial a_2}{\partial x_1} - \frac{\partial a_1}{\partial x_2} \right).$$

Note that if we had let the Levi–Civita tensor act on the transpose tensor $(\nabla \boldsymbol{a})^T$ we obtain

$${}_3\boldsymbol{\epsilon} : (\nabla \boldsymbol{a})^T = \nabla \times \boldsymbol{a}, \tag{1.80}$$

which shows that

$${}_3\boldsymbol{\epsilon} : \left[\nabla \boldsymbol{a} + (\nabla \boldsymbol{a})^T \right] = \nabla \times \boldsymbol{a} - \nabla \times \boldsymbol{a} = 0. \tag{1.81}$$

This fact will be used later in the proof of *Curie's principle* as given in Section 1.8.5.

If the curl operator acts on a second-order tensor field $\boldsymbol{A} = A_{mn} \hat{\boldsymbol{x}}_m \hat{\boldsymbol{x}}_n$, the result is a second-order tensor

$$\nabla \times \boldsymbol{A} = -{}_3\boldsymbol{\epsilon} : \nabla \boldsymbol{A} = -\epsilon_{ijk} \frac{\partial A_{jn}}{\partial x_k} \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_n \tag{1.82}$$

$$= \left[\hat{\boldsymbol{x}}_1 \left(\frac{\partial A_{3n}}{\partial x_2} - \frac{\partial A_{2n}}{\partial x_3} \right) + \hat{\boldsymbol{x}}_2 \left(\frac{\partial A_{1n}}{\partial x_3} - \frac{\partial A_{3n}}{\partial x_1} \right) + \hat{\boldsymbol{x}}_3 \left(\frac{\partial A_{2n}}{\partial x_1} - \frac{\partial A_{1n}}{\partial x_2} \right) \right] \hat{\boldsymbol{x}}_n. \tag{1.83}$$

Being able to do curl operations using the dot product makes it possible to prove many useful things about curl operations.

For example, in both fluid mechanics and electromagnetism, frequent use is made of the identity

$$\nabla \times \nabla \times \boldsymbol{u} = \nabla(\nabla \cdot \boldsymbol{u}) - \nabla^2 \boldsymbol{u}. \tag{1.84}$$

We can prove this identity using the Levi–Civita tensor as follows:

$$\nabla \times \nabla \times \boldsymbol{u} = \epsilon_{ijk}\, \hat{\boldsymbol{x}}_i \hat{\boldsymbol{x}}_j \hat{\boldsymbol{x}}_k : \frac{\partial}{\partial x_l} \hat{\boldsymbol{x}}_l \left[\epsilon_{mno}\, \hat{\boldsymbol{x}}_m \hat{\boldsymbol{x}}_n \hat{\boldsymbol{x}}_o : \hat{\boldsymbol{x}}_p \frac{\partial}{\partial x_p} u_q \hat{\boldsymbol{x}}_q \right] \tag{1.85}$$
$$= \epsilon_{ijk}\, \epsilon_{mno} \frac{\partial^2 u_q}{\partial x_l\, \partial x_p} \hat{\boldsymbol{x}}_i \left(\hat{\boldsymbol{x}}_k \cdot \hat{\boldsymbol{x}}_l\right) \left(\hat{\boldsymbol{x}}_j \cdot \hat{\boldsymbol{x}}_m\right) \left(\hat{\boldsymbol{x}}_o \cdot \hat{\boldsymbol{x}}_p\right) \left(\hat{\boldsymbol{x}}_n \cdot \hat{\boldsymbol{x}}_q\right).$$

So $l = k$, $m = j$, $p = o$, and $q = n$, to give

$$\nabla \times \nabla \times \boldsymbol{u} = \epsilon_{ijk}\, \epsilon_{jno} \frac{\partial^2 u_n}{\partial x_k\, \partial x_o} \hat{\boldsymbol{x}}_i. \tag{1.86}$$

As shown earlier, we have the identity that $\epsilon_{ijk}\epsilon_{ino} = \delta_{jn}\delta_{ko} - \delta_{jo}\delta_{kn}$. If we exchange the indices i and j and use that $\epsilon_{jik} = -\epsilon_{ijk}$ we then have the identity

$$\epsilon_{ijk}\, \epsilon_{jno} = \delta_{io}\delta_{kn} - \delta_{in}\delta_{ko}. \tag{1.87}$$

Using this in Eq. (1.86) gives the sought after result

$$\nabla \times \nabla \times \boldsymbol{u} = \hat{\boldsymbol{x}}_i \frac{\partial}{\partial x_i} \left(\frac{\partial u_k}{\partial x_k} \right) - \frac{\partial^2 u_i}{\partial x_k^2} \hat{\boldsymbol{x}}_i \tag{1.88}$$
$$= \nabla(\nabla \cdot \boldsymbol{u}) - \nabla^2 \boldsymbol{u}. \tag{1.89}$$

Using ${}_3\boldsymbol{\epsilon}$, it is also straightforward to show

$$\nabla \times \nabla \alpha = 0 \tag{1.90}$$
$$\nabla \cdot (\nabla \times \boldsymbol{a}) = 0 \tag{1.91}$$

both of which are used throughout continuum physics.

1.7 Tensor-Calculus Product Rules

What if the gradient operator ∇ in an expression acts on several vectorial or tensorial terms with, possibly, various scalar products present between the vectors and tensors? What is the equivalent of the scalar "product rule" $\partial(\alpha\beta)/\partial x = \beta\, \partial\alpha/\partial x + \alpha\, \partial\beta/\partial x$ for various types of products involving vectors and tensors?

As an example, let's consider the specific expression

$$\nabla \cdot (\boldsymbol{ab} \cdot \boldsymbol{T}), \tag{1.92}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are both vectors and $\boldsymbol{T}$ is a second-order tensor, all of which vary in space. In this expression, $\boldsymbol{ab}$ is a second-order tensor and so is $\boldsymbol{ab} \cdot \boldsymbol{T}$ so that $\nabla \cdot (\boldsymbol{ab} \cdot \boldsymbol{T})$ is a

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

vector. To distribute the derivative in this expression, we work in Cartesian coordinates in which the base vectors $\hat{\boldsymbol{x}}_i$ are uniform so that $\partial\hat{\boldsymbol{x}}_j/\partial x_i = 0$ for all base vectors and coordinate directions. But after we distribute the derivatives in Cartesians, we will write the resulting expression in the general bold-face notation that then applies to even curvilinear coordinates in which some of the base vectors can have nonzero derivatives. Any tensor identity expressed in bold-face notation is valid for any coordinate system and is the preferred way to express tensor-calculus product rules.

For each vectorial or tensorial term in the expression, we use a different set of Cartesian-coordinate indices to write out

$$\nabla\cdot(\boldsymbol{ab}\cdot\boldsymbol{T}) = \underbrace{\hat{\boldsymbol{x}}_i\frac{\partial}{\partial x_i}}_{\nabla}\cdot\left(\underbrace{a_j\hat{\boldsymbol{x}}_j\, b_k\hat{\boldsymbol{x}}_k}_{\boldsymbol{ab}}\cdot\underbrace{T_{lm}\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m}_{\boldsymbol{T}}\right) \tag{1.93}$$

$$= \frac{\partial}{\partial x_i}\left(a_j b_k T_{lm}\right)\left(\hat{\boldsymbol{x}}_i\cdot\hat{\boldsymbol{x}}_j\right)\left(\hat{\boldsymbol{x}}_k\cdot\hat{\boldsymbol{x}}_l\right)\,\hat{\boldsymbol{x}}_m. \tag{1.94}$$

In passing from the first line to the second, we pulled out the scalar components and derivative while preserving the position of the base vectors and scalar products between the base vectors. Next, because of the nature of the scalar product in orthogonal coordinates, we have that $j = i$ and $l = k$. This allows us to write

$$\nabla\cdot(\boldsymbol{ab}\cdot\boldsymbol{T}) = \frac{\partial}{\partial x_i}\left(a_i b_k T_{km}\right)\hat{\boldsymbol{x}}_m \tag{1.95}$$

$$= \left[\left(\frac{\partial a_i}{\partial x_i}\right) b_k T_{km} + a_i\left(\frac{\partial b_k}{\partial x_i}\right) T_{km} + a_i b_k \frac{\partial T_{km}}{\partial x_i}\right]\hat{\boldsymbol{x}}_m. \tag{1.96}$$

In going from the first to second expression, we just employed the usual derivative product rule for scalar fields.

The final step is what students often find the most difficult. One must look at Eq. (1.96) and identify the equivalent expression in bold face. So you have to make identifications like $\partial a_i/\partial x_i = \nabla\cdot\boldsymbol{a}$ and $b_k T_{km}\hat{\boldsymbol{x}}_m = \boldsymbol{b}\cdot\boldsymbol{T}$. Carrying this out, we obtain at last

$$\nabla\cdot(\boldsymbol{ab}\cdot\boldsymbol{T}) = (\nabla\cdot\boldsymbol{a})\,\boldsymbol{b}\cdot\boldsymbol{T} + \boldsymbol{a}\cdot(\nabla\boldsymbol{b})\cdot\boldsymbol{T} + \boldsymbol{b}\cdot(\boldsymbol{a}\cdot\nabla\boldsymbol{T})\,. \tag{1.97}$$

Once we type the final expression in bold face (or write by hand the expression with squiggly underscores), it applies to any orthogonal curvilinear coordinates and is a generally valid identity not limited to Cartesian coordinates. Note that $\nabla\boldsymbol{T}$ is an example of a third-order tensor. We have thus determined how to distribute the ∇ operator onto the vectors and tensors in a multi-term expression involving scalar products. By using the earlier Levi–Civita alternating tensor, we can do the same for multi-term expressions involving vector products.

Using this ability, it is now straightforward to derive a long list of useful tensor-calculus product rules. In the following list, α is a scalar field, $\boldsymbol{a}$ and $\boldsymbol{b}$ are again vector fields, $\boldsymbol{A}$ and $\boldsymbol{B}$ are second-order tensor fields and ${}_4\boldsymbol{C}$ is a fourth-order tensor field:

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\nabla \times (\alpha \boldsymbol{a}) = \alpha \nabla \times \boldsymbol{a} + (\nabla \alpha) \times \boldsymbol{a} \tag{1.98}$$

$$\nabla \times (\boldsymbol{a} \times \boldsymbol{b}) = \nabla \cdot (\boldsymbol{ba} - \boldsymbol{ab}) \tag{1.99}$$

$$\nabla \cdot (\boldsymbol{ab}) = (\nabla \cdot \boldsymbol{a})\boldsymbol{b} + \boldsymbol{a} \cdot \nabla \boldsymbol{b} \tag{1.100}$$

$$\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{b} \cdot (\nabla \times \boldsymbol{a}) - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b}) \tag{1.101}$$

$$\nabla \cdot (\boldsymbol{AB}) = (\nabla \cdot \boldsymbol{A})\boldsymbol{B} + \boldsymbol{A}^T \cdot \nabla \boldsymbol{B} \tag{1.102}$$

$$\nabla \cdot (\boldsymbol{A} \cdot \boldsymbol{B}) = (\nabla \cdot \boldsymbol{A}) \cdot \boldsymbol{B} + \boldsymbol{A}^T : \nabla \boldsymbol{B} \tag{1.103}$$

$$\nabla (\boldsymbol{a} \cdot \boldsymbol{b}) = (\nabla \boldsymbol{a}) \cdot \boldsymbol{b} + (\nabla \boldsymbol{b}) \cdot \boldsymbol{a} \tag{1.104}$$

$$= \boldsymbol{a} \cdot (\nabla \boldsymbol{b}) + \boldsymbol{b} \cdot (\nabla \boldsymbol{a}) + \boldsymbol{a} \times (\nabla \times \boldsymbol{b}) + \boldsymbol{b} \times (\nabla \times \boldsymbol{a}) \tag{1.105}$$

$$\nabla[(\nabla \alpha) \cdot \nabla \alpha] = 2(\nabla \alpha) \cdot \nabla \nabla \alpha = 2(\nabla \nabla \alpha) \cdot \nabla \alpha \tag{1.106}$$

$$\nabla (\boldsymbol{A} \cdot \boldsymbol{B}) = (\nabla \boldsymbol{A}) \cdot \boldsymbol{B} + [(\nabla \boldsymbol{B}) \cdot \boldsymbol{A}]^{\overset{T}{132}} \tag{1.107}$$

$$\nabla (\boldsymbol{A} \cdot \boldsymbol{a}) = (\nabla \boldsymbol{A}) \cdot \boldsymbol{a} + (\nabla \boldsymbol{a}) \cdot \boldsymbol{A}^T \tag{1.108}$$

$$\nabla (\boldsymbol{a} \cdot \boldsymbol{A}) = \left[\nabla \left(\boldsymbol{A}^T\right)\right] \cdot \boldsymbol{a} + (\nabla \boldsymbol{a}) \cdot \boldsymbol{A} \tag{1.109}$$

$$\nabla \cdot (\alpha \boldsymbol{A}) = \nabla \alpha \cdot \boldsymbol{A} + \alpha \nabla \cdot \boldsymbol{A} \tag{1.110}$$

$$\nabla \cdot (\boldsymbol{a} \cdot \boldsymbol{A}) = \nabla \boldsymbol{a} : \boldsymbol{A} + \left(\nabla \cdot \boldsymbol{A}^T\right) \cdot \boldsymbol{a} \tag{1.111}$$

$$\nabla \cdot (\boldsymbol{A} \cdot \boldsymbol{a}) = \boldsymbol{A} : (\nabla \boldsymbol{a})^T + (\nabla \cdot \boldsymbol{A}) \cdot \boldsymbol{a} \tag{1.112}$$

$$\nabla \cdot (\boldsymbol{aA}) = (\nabla \cdot \boldsymbol{a})\boldsymbol{A} + \boldsymbol{a} \cdot \nabla \boldsymbol{A} \tag{1.113}$$

$$\nabla \cdot (\boldsymbol{Aa}) = (\nabla \cdot \boldsymbol{A})\, \boldsymbol{a} + \boldsymbol{A}^T \cdot \nabla \boldsymbol{a} \tag{1.114}$$

$$\nabla \cdot ({}_4\boldsymbol{C} : \boldsymbol{A}) = (\nabla \cdot {}_4\boldsymbol{C}) : \boldsymbol{A} + {}_4\boldsymbol{C}^{\overset{T}{2341}} \overset{3}{\vdots} \nabla \boldsymbol{A} \tag{1.115}$$

Note that in the last identity, the symbol $\overset{3}{\vdots}$ means “the triple dot product” as defined earlier. Having the above types of tensor-calculus product rules available to us for arbitrary curvilinear coordinates allows many results in continuum physics to be developed in the chapters that follow.

It may not seem obvious that tensor-calculus product rules proven in Cartesian coordinates, in which derivatives of the base vectors are zero, but written in bold-face notation after distributing the derivatives are in fact generally valid for all orthogonal curvilinear coordinate systems. To demonstrate this fact using an example, consider the identity $\nabla \cdot (\boldsymbol{A} \cdot \boldsymbol{a}) = \boldsymbol{A} : (\nabla \boldsymbol{a})^T + (\nabla \cdot \boldsymbol{A}) \cdot \boldsymbol{a}$ given in the above list as derived, for convenience, in Cartesian coordinates. To prove this identity is valid in say cylindrical coordinates, we simply carry out all the given operations using that $\partial \hat{\boldsymbol{r}}/\partial \phi = \hat{\boldsymbol{\phi}}$ and $\partial \hat{\boldsymbol{\phi}}/\partial \phi = -\hat{\boldsymbol{r}}$ with all other derivatives of the base vectors equal to zero. In cylindrical coordinates, the various terms are

$$\nabla \boldsymbol{a} = \begin{pmatrix} \partial a_r/\partial r & \partial a_\phi/\partial r & \partial a_z/\partial r \\ (\partial a_r/\partial \phi - a_\phi)/r & (\partial a_\phi/\partial \phi + a_r)/r & \partial a_z/\partial \phi \\ \partial a_r/\partial z & \partial a_\phi/\partial z & \partial a_z/\partial z \end{pmatrix}, \tag{1.116}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

a result that will be obtained more formally in the upcoming Section 1.8.6 on orthogonal curvilinear coordinates. We then have that

$$\begin{aligned}\boldsymbol{A}:(\nabla\boldsymbol{a})^T &= \operatorname{tr}\left\{\boldsymbol{A}\cdot(\nabla\boldsymbol{a})^T\right\}\\ &= A_{rr}\frac{\partial a_r}{\partial r}+A_{r\phi}\frac{\partial a_\phi}{\partial r}+A_{rz}\frac{\partial a_z}{\partial r}\\ &\quad+\frac{A_{\phi r}}{r}\left(\frac{\partial a_r}{\partial\phi}-a_\phi\right)+\frac{A_{\phi\phi}}{r}\left(\frac{\partial a_\phi}{\partial\phi}+a_r\right)+A_{\phi z}\frac{\partial a_z}{\partial\phi}\\ &\quad+A_{zr}\frac{\partial a_r}{\partial z}+A_{z\phi}\frac{\partial a_\phi}{\partial z}+A_{zz}\frac{\partial a_z}{\partial z},\end{aligned} \tag{1.117}$$

$$\begin{aligned}\nabla\cdot\boldsymbol{A} &= \hat{\boldsymbol{r}}\left(\frac{1}{r}\frac{\partial(rA_{rr})}{\partial r}+\frac{1}{r}\frac{\partial A_{\phi r}}{\partial\phi}-\frac{A_{\phi\phi}}{r}+\frac{A_{zr}}{\partial z}\right)\\ &\quad+\hat{\boldsymbol{\phi}}\left(\frac{1}{r}\frac{\partial(rA_{r\phi})}{\partial r}+\frac{1}{r}\frac{\partial A_{\phi\phi}}{\partial\phi}+\frac{A_{\phi r}}{r}+\frac{A_{z\phi}}{\partial z}\right)\\ &\quad+\hat{\boldsymbol{z}}\left(\frac{1}{r}\frac{\partial(rA_{rz})}{\partial r}+\frac{1}{r}\frac{\partial A_{\phi z}}{\partial\phi}+\frac{A_{zz}}{\partial z}\right),\end{aligned} \tag{1.118}$$

and

$$\begin{aligned}\nabla\cdot(\boldsymbol{A}\cdot\boldsymbol{a}) &= \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(A_{rr}a_r+A_{r\phi}a_\phi+A_{rz}a_z\right)\right]\\ &\quad+\frac{1}{r}\frac{\partial}{\partial\phi}\left(A_{\phi r}a_r+A_{\phi\phi}a_\phi+A_{\phi z}a_z\right)\\ &\quad+\frac{\partial}{\partial z}\left(A_{zr}a_r+A_{z\phi}a_\phi+A_{zz}a_z\right).\end{aligned} \tag{1.119}$$

Using these expressions, a final bit of algebra demonstrates that $\nabla\cdot(\boldsymbol{A}\cdot\boldsymbol{a})=\boldsymbol{A}:(\nabla\boldsymbol{a})^T+(\nabla\cdot\boldsymbol{A})\cdot\boldsymbol{a}$ is valid in cylindrical coordinates despite having been derived initially in Cartesian coordinates.

1.8 Additional Topics Involving Tensors

Each topic treated in this section is important because it will be used in our development of the rules of continuum physics. However, at this point, you have been exposed to enough tensor calculus that after first working through Section 1.9 on the Dirac delta function and working some end-of-chapter exercises to sharpen your skills, you can move ahead to Chapter 2 on continuum mechanics if you so choose. The topics treated in this section will be referred to each time they are needed in the chapters to follow. On the other hand, working through this section now will make you more proficient with tensors and tensor calculus, will show you some interesting uses and facts of tensors, and will better prepare you for the chapters that follow.

1.8.1 Taylor Series of Fields in Three-Dimensional Space

It is useful in various physics contexts to represent a scalar, vector, or tensor field near a particular point as a power series in the local coordinates adjacent to that point if the field and its spatial derivatives are known at that point. This is what the Taylor-series expansion does for us and, in three-dimensional space, requires the use of higher-order tensors and dot products, even for the representation of scalar fields.

But let's begin with the well-known example of a scalar function $\psi(x)$ in just one spatial dimension x that does not require the use of tensors. We expand $\psi(x)$ about a particular point x_0 as a power series that is called the Taylor series

$$\boxed{\psi(x) = \sum_{n=0}^{\infty} (x - x_0)^n a_n.} \tag{1.120}$$

Because $(x - x_0)^0 = 1$, the first coefficient a_0 is found by simply evaluating this series at $x = x_0$

$$\psi(x_0) = a_0. \tag{1.121}$$

Each successive coefficient is found by taking successive derivatives and evaluating at $x = x_0$

$$\left.\frac{\mathrm{d}\psi}{\mathrm{d}x}\right|_{x_0} = a_1, \quad \left.\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2}\right|_{x_0} = (2)(1)a_2, \quad \text{and} \quad \left.\frac{\mathrm{d}^3\psi}{\mathrm{d}x^3}\right|_{x_0} = (3)(2)(1)a_3 \tag{1.122}$$

so that each coefficient in the Taylor series is given by the derivatives of the function at the point x_0 as

$$\boxed{a_n = \frac{1}{n!} \left.\frac{\mathrm{d}^n\psi}{\mathrm{d}x^n}\right|_{x_0} \quad \text{for} \quad n = 1, 2 \ldots \infty} \tag{1.123}$$

and where, again, $a_0 = \psi(x_0)$.

Next, for the Taylor series of any field (scalar, vector, or tensor) that is distributed in three-dimensional space, we again expand this field in a power series of the local coordinates about a particular point $\boldsymbol{r}_0$ but now each term n in the series involves coefficients that are tensors of (at least) order n. For a field written as ${}_t\boldsymbol{\psi}(\boldsymbol{r})$, where t denotes the tensorial order of the field (so $t = 0$ is a scalar field, $t = 1$ a vector field and so on), the Taylor series expansion for this field is written

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$ {}_t\boldsymbol{\psi}(\boldsymbol{r}) = \sum_{n=0}^{\infty} (\boldsymbol{r}-\boldsymbol{r}_0)^n \overset{n}{:} {}_{\{n+t\}}\boldsymbol{A}, \tag{1.124}$$

where the notation is made clear by writing out the first few terms of the series

$$\begin{aligned} {}_t\boldsymbol{\psi}(\boldsymbol{r}) = {}_t\boldsymbol{A} &+ (\boldsymbol{r}-\boldsymbol{r}_0)\cdot {}_{\{1+t\}}\boldsymbol{A} \\ &+ (\boldsymbol{r}-\boldsymbol{r}_0)(\boldsymbol{r}-\boldsymbol{r}_0) : {}_{\{2+t\}}\boldsymbol{A} \\ &+ (\boldsymbol{r}-\boldsymbol{r}_0)(\boldsymbol{r}-\boldsymbol{r}_0)(\boldsymbol{r}-\boldsymbol{r}_0) \overset{3}{:} {}_{\{3+t\}}\boldsymbol{A} + \dots . \end{aligned} \tag{1.125}$$

So for a field of tensorial order t in three-dimensional space, the Taylor-series coefficients at each n are tensors ${}_{\{n+t\}}\boldsymbol{A}$ of tensorial order $n+t$ (i.e., that have $n+t$ base vectors).

To find these tensorial coefficients, begin by evaluating the tensorial power series at $\boldsymbol{r}=\boldsymbol{r}_0$ to give the first tensorial coefficient as

$$ {}_t\boldsymbol{\psi}(\boldsymbol{r}_0) = {}_t\boldsymbol{A}. \tag{1.126}$$

The subsequent tensorial coefficients are obtained by taking successive gradients of the series and evaluating at $\boldsymbol{r}=\boldsymbol{r}_0$ beginning with

$$ \nabla\ ({}_t\boldsymbol{\psi})|_{\boldsymbol{r}_0} = \nabla\boldsymbol{r}\cdot {}_{\{t+1\}}\boldsymbol{A} = {}_{\{1+t\}}\boldsymbol{A}, \tag{1.127}$$

$$ \nabla\nabla\ ({}_t\boldsymbol{\psi})|_{\boldsymbol{r}_0} = \nabla\nabla\,(\boldsymbol{r}\boldsymbol{r}) : {}_{\{2+t\}}\boldsymbol{A}, \tag{1.128}$$

where the tensorial coefficients ${}_{\{n+t\}}\boldsymbol{A}$ are constants (the ∇ acting on them gives zero) and where in Eq. (1.127) we used the earlier result that $\nabla\boldsymbol{r}=\boldsymbol{I}$ is the second-order identity tensor. It is a straightforward exercise to show that $\nabla\nabla\,(\boldsymbol{r}\boldsymbol{r}) = \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ with summation over the repeated indices as always, which is a type of fourth-order identity-transpose tensor. To work with it, we double dot it into the tensorial coefficient ${}_{\{2+t\}}\boldsymbol{A} = A_{kl\dots\alpha}\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\dots\hat{\boldsymbol{x}}_\alpha$, where how many base vectors this tensorial coefficient has beyond the first two depends on the tensorial order t of the field being expanded. A scalar field ($t=0$) has no additional base vectors, a vector field ($t=1$) one additional base vector and so on. We have

$$ \nabla\nabla\,(\boldsymbol{r}\boldsymbol{r}) : {}_{\{2+t\}}\boldsymbol{A} = \left(\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\right) : A_{kl\dots\alpha}\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\dots\hat{\boldsymbol{x}}_\alpha \tag{1.129}$$

$$ = \left(A_{ij\dots\alpha} + A_{ji\dots\alpha}\right)\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\dots\hat{\boldsymbol{x}}_\alpha . \tag{1.130}$$

Because $\nabla\nabla(\boldsymbol{r}\boldsymbol{r})$ is symmetric in the last two base vectors, Eq. (1.128) requires the tensorial coefficient ${}_{\{2+t\}}\boldsymbol{A}$ to be symmetric in the first two base vectors, so that

$$ \nabla\nabla\ ({}_t\boldsymbol{\psi})|_{\boldsymbol{r}_0} = (2)\,{}_{\{2+t\}}\boldsymbol{A}. \tag{1.131}$$

An identical analysis for the third-order coefficient that exploits the symmetry of $\nabla\nabla\nabla(\boldsymbol{r}\boldsymbol{r}\boldsymbol{r})$ yields

$$ \nabla\nabla\nabla\ ({}_t\boldsymbol{\psi})|_{\boldsymbol{r}_0} = (3)(2)\,{}_{\{3+t\}}\boldsymbol{A} \tag{1.132}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

with ${}_{\{3+t\}}\boldsymbol{A}$ having complete symmetry between the first three base vectors. Thus, the nth tensorial coefficient in the Taylor series of a tensorial field is given as

$$
{}_{\{n+t\}}\boldsymbol{A} = \frac{1}{n!}\, {}_n\nabla \left({}_t\boldsymbol{\psi}\right)\big|_{\boldsymbol{r}_0} \quad \text{for} \quad n = 1, 2 \ldots \infty, \tag{1.133}
$$

where ${}_n\nabla = \nabla\nabla \ldots \nabla$ means n successive applications of the gradient operator acting on the tensor field ${}_t\boldsymbol{\psi}$ of order t before evaluating at the particular point in 3D space $\boldsymbol{r} = \boldsymbol{r}_0$. The leading $n = 0$ coefficient is again ${}_t\boldsymbol{A} = {}_t\boldsymbol{\psi}(\boldsymbol{r}_0)$.

So as an example from Chapter 2, we may wish to represent the electric field $\boldsymbol{E}(\boldsymbol{r})$ within a molecule whose center is located at $\boldsymbol{r}_0$ as an explicit function in the local coordinates $\boldsymbol{r} - \boldsymbol{r}_0$ within the molecule. We thus perform a Taylor-series expansion of the electric field about $\boldsymbol{r}_0$ to give

$$
\begin{aligned}
\boldsymbol{E}(\boldsymbol{r}) = {} & \boldsymbol{E}(\boldsymbol{r}_0) + (\boldsymbol{r} - \boldsymbol{r}_0) \cdot \nabla \boldsymbol{E}|_{\boldsymbol{r}_0} + \\
& + \frac{1}{2!}\,(\boldsymbol{r} - \boldsymbol{r}_0)\,(\boldsymbol{r} - \boldsymbol{r}_0) : \nabla\nabla\boldsymbol{E}|_{\boldsymbol{r}_0} + O\left(|\boldsymbol{r} - \boldsymbol{r}_0|^3\right).
\end{aligned} \tag{1.134}
$$

The notation $O\left(|\boldsymbol{r} - \boldsymbol{r}_0|^3\right)$ is called the "big-O" notation and is used to represent the part of the series that is being truncated in a certain limit such as $\boldsymbol{r} - \boldsymbol{r}_0 \to 0$. In this limit for this particular example, the amplitude $|\boldsymbol{r} - \boldsymbol{r}_0|^3$ is the largest part of what is being truncated in Eq. (1.134), which is what the notation $O\left(|\boldsymbol{r} - \boldsymbol{r}_0|^3\right)$ is saying. Let's write a power series in the parameter ϵ as $f(\epsilon) = a_0 + a_1\epsilon + a_2\epsilon^2 + a_3\epsilon^3 + \ldots$. If in the limit of $\epsilon \to 0$, we truncate the series after the first two terms, we use the big-O notation to write $f(\epsilon) = a_0 + a_1\epsilon + O(\epsilon^2)$, where the argument of $O(\epsilon^2)$ represents the size or "order" of what is being neglected in the limit, which in this case is $O(\epsilon^2) = a_2\epsilon^2[1 + (a_3/a_2)\epsilon + (a_4/a_2)\epsilon^2 + \ldots] \to a_2\epsilon^2$ as $\epsilon \to 0$.

1.8.2 Functions of Second-Order Tensors

It will arise that we want to consider a function whose argument is a second-order tensor, that is, $f(\boldsymbol{A})$ where both $\boldsymbol{A}$ and $f(\boldsymbol{A})$ are second-order tensors. What do we mean by this?

If $f(\alpha)$ is some function of a scalar α, we expand $f(\alpha)$ as a Taylor series about $\alpha = 0$ as

$$
\begin{aligned}
f(\alpha) = {} & f(0) + \frac{1}{1!}\left.\frac{\mathrm{d}f(\alpha)}{\mathrm{d}\alpha}\right|_{\alpha=0} \alpha \\
& + \frac{1}{2!}\left.\frac{\mathrm{d}^2 f(\alpha)}{\mathrm{d}\alpha^2}\right|_{\alpha=0} \alpha^2 + \frac{1}{3!}\left.\frac{\mathrm{d}^3 f(\alpha)}{\mathrm{d}\alpha^3}\right|_{\alpha=0} \alpha^3 + \ldots.
\end{aligned} \tag{1.135}
$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

Because α is a scalar, so is $f(\alpha)$. We now define the second-order tensor $f(\boldsymbol{A})$ as the operation or rule

$$f(\boldsymbol{A}) = f(0)\boldsymbol{I} + \frac{1}{1!}\left.\frac{\mathrm{d}f(\alpha)}{\mathrm{d}\alpha}\right|_{\alpha=0}\boldsymbol{A} + \frac{1}{2!}\left.\frac{\mathrm{d}^2 f(\alpha)}{\mathrm{d}\alpha^2}\right|_{\alpha=0}\boldsymbol{A}\cdot\boldsymbol{A} + \frac{1}{3!}\left.\frac{\mathrm{d}^3 f(\alpha)}{\mathrm{d}\alpha^3}\right|_{\alpha=0}\boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots, \tag{1.136}$$

where $\boldsymbol{A}$ is a second-order tensor, $\boldsymbol{I}$ is the second-order identity tensor, and each term in the series is a second-order tensor.

Let's give some examples. Consider the specific scalar function $f_1(\alpha) = (1-\alpha)^{-1} = 1 + \alpha + \alpha^2 + \ldots$. We then can define the second-order tensor $f_1(\boldsymbol{A})$ operation as $f_1(\boldsymbol{A}) = (\boldsymbol{I} - \boldsymbol{A})^{-1}$, where $\boldsymbol{I}$ is again the second-order identity tensor. The operation $(\boldsymbol{I}-\boldsymbol{A})^{-1}$ is understood, through Eq. (1.136), to mean the expansion

$$(\boldsymbol{I}-\boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots. \tag{1.137}$$

We then expect that $(\boldsymbol{I}-\boldsymbol{A})\cdot(\boldsymbol{I}-\boldsymbol{A})^{-1} = \boldsymbol{I}$, which can be verified through explicit multiplication

$$(\boldsymbol{I}-\boldsymbol{A})\cdot(\boldsymbol{I}-\boldsymbol{A})^{-1} = (\boldsymbol{I}-\boldsymbol{A})\cdot(\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots) = \boldsymbol{I}. \tag{1.138}$$

So a function of a second-order tensor, that is itself a second-order tensor, is coherently defined through the expansion of Eq. (1.136).

As another specific example, define a second-order tensor $\boldsymbol{B}$ as the function of another second-order tensor $\boldsymbol{A}$ through the operations

$$\boldsymbol{B} = -\ln(\boldsymbol{I}-\boldsymbol{A}) = \boldsymbol{A} + \frac{1}{2}\boldsymbol{A}\cdot\boldsymbol{A} + \frac{1}{3}\boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots. \tag{1.139}$$

We can then take the exponential of $-\boldsymbol{B}$ by which we mean the expansion

$$\begin{aligned}\exp(-\boldsymbol{B}) &= \boldsymbol{I} - \frac{1}{1!}\boldsymbol{B} + \frac{1}{2!}\boldsymbol{B}\cdot\boldsymbol{B} - \frac{1}{3!}\boldsymbol{B}\cdot\boldsymbol{B}\cdot\boldsymbol{B} + \ldots \\ &= \exp(\ln(\boldsymbol{I}-\boldsymbol{A})) = \boldsymbol{I}-\boldsymbol{A}.\end{aligned} \tag{1.140}$$

This last relation then gives

$$\boldsymbol{A} = \boldsymbol{I} - \exp(-\boldsymbol{B}) = \frac{1}{1!}\boldsymbol{B} - \frac{1}{2!}\boldsymbol{B}\cdot\boldsymbol{B} + \frac{1}{3!}\boldsymbol{B}\cdot\boldsymbol{B}\cdot\boldsymbol{B} + \ldots. \tag{1.141}$$

If we then substitute the original definition of $\boldsymbol{B}$ from Eq. (1.139), we obtain

$$\boldsymbol{A} = \frac{1}{1!}\left(\boldsymbol{A} + \frac{1}{2}\boldsymbol{A}\cdot\boldsymbol{A} + \frac{1}{3}\boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots\right) - \frac{1}{2!}(\boldsymbol{A}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots) + \frac{1}{3!}(\boldsymbol{A}\cdot\boldsymbol{A}\cdot\boldsymbol{A} + \ldots) + \ldots \tag{1.142}$$

$$= \boldsymbol{A}. \tag{1.143}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

So these operations are internally consistent.

To conclude, the function of a second-order tensor is defined here to be another second-tensor tensor as calculated by the expansion of Eq. (1.136). We can also define the inverse function of a second-order tensor as we have shown in the above examples that we will see again later in the book.

1.8.3 Rotation of the Cartesian Coordinates

As we have emphasized, a tensor of any order, including a first-order tensor or vector, is a field that exists at each point in space independently of whatever coordinates we choose to work in. But it can arise that we want to work in a Cartesian coordinate system $\hat{\boldsymbol{x}}'_i$ that has been rotated from an initial system $\hat{\boldsymbol{x}}_i$ as shown in Fig. 1.12. We would like to know how the *scalar components* of the vectors and tensors change when we rotate the base vectors to have new orientations. It is sometimes stated that a tensor is defined by the rules derived below for how the Cartesian components of the tensor change with the changing orientation of the base vectors. However, we have already seen that tensors of any order are coherently defined without first having in place such "coordinate-rotation rules." The fundamental nature (and need) of tensors as used in continuum physics is again that they map, using dot products, one tensor (including vectors) into another tensor and such tensorial mappings exist independently of knowing how the Cartesian coefficients of a tensor change with the orientation of the coordinates.

A rotation of angle θ_1 about the $\hat{\boldsymbol{x}}_1$ axis is allowed for by the matrix operation (θ_1 positive is in a counterclockwise sense when $\hat{\boldsymbol{x}}_1$ is oriented toward the observer)

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \tag{1.144}$$

which is easily confirmed by doing the trigonometry in Fig. 1.13.

We define the "rotation matrix" for rotations around the x_1 axis as

$$R^{(1)}_{ij}(\theta_1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix}. \tag{1.145}$$

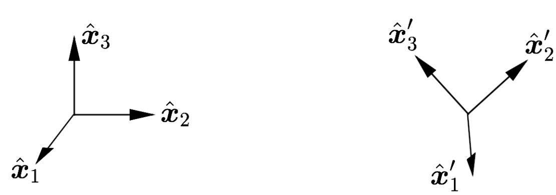

Figure 1.12 Two Cartesian-coordinate systems that are rotated relative to each other.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

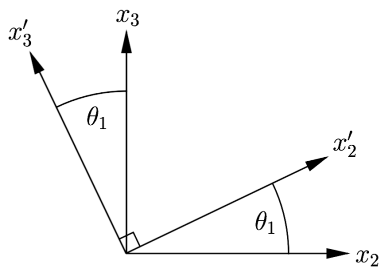

Figure 1.13 Rotating the coordinates counterclockwise by an angle θ_1 around the x_1 axis.

The components of a vector $\boldsymbol{a} = a_1\,\hat{\boldsymbol{x}}_1 + a_2\,\hat{\boldsymbol{x}}_2 + a_3\,\hat{\boldsymbol{x}}_3$ transform in the rotated coordinate system to $a_i' = R_{ij}^{(1)}(\theta_1)\,a_j$ with summation over the repeated index j being performed as the operation

$$\begin{bmatrix} a_1' \\ a_2' \\ a_3' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}. \tag{1.146}$$

We just have to carry out the matrix multiplication.

Similarly, a second-order tensor $\boldsymbol{T} = T_{ij}\,\hat{\boldsymbol{x}}_i\,\hat{\boldsymbol{x}}_j$ has Cartesian components T_{ij} that transform as

$$T_{ij}' = R_{ik}^{(1)}(\theta_1)\,R_{jl}^{(1)}(\theta_1)\,T_{kl} \quad \text{(sum over repeated indices)}, \tag{1.147}$$

where we have to apply the rotation matrix to each base vector. If we write the sums over repeated indices using matrix multiplication (the inner product), we must rearrange this expression as $T_{ij}' = R_{ik}^{(1)} T_{kl} \left(R_{jl}^{(1)}\right)^T$ so that the position of the indices correspond to the inner product and the matrix operation

$$\begin{bmatrix} T_{11}' & T_{12}' & T_{13}' \\ T_{21}' & T_{22}' & T_{23}' \\ T_{31}' & T_{32}' & T_{33}' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{bmatrix}. \tag{1.148}$$

Note that if $\boldsymbol{T}$ is proportional to the second-order identity tensor $\boldsymbol{T} = T\delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$, we have

$$\begin{bmatrix} T_{11}' & T_{12}' & T_{13}' \\ T_{21}' & T_{22}' & T_{23}' \\ T_{31}' & T_{32}' & T_{33}' \end{bmatrix} = T \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{bmatrix} \tag{1.149}$$

$$= T \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.150}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

So the second-order identity tensor satisfies $T'_{ij} = T\delta'_{ij} = T_{ij} = T\delta_{ij}$ or $\delta'_{ij} = \delta_{ij}$ for any rotation of the coordinates about the x_1 axis.

Similarly, a fourth-order tensor has components that transform with coordinate rotations about the x_1 axis as

$$C'_{ijkl} = R^{(1)}_{im}(\theta_1)\, R^{(1)}_{jn}(\theta_1)\, R^{(1)}_{ko}(\theta_1)\, R^{(1)}_{lp}(\theta_1)\, C_{mnop} \tag{1.151}$$

with summation over the repeated indices m, n, o, and p. Higher-order tensors are handled in an analogous manner, using one rotation matrix for each base vector.

A rotation about the x_2 axis is accomplished using (θ_2 is again in the counterclockwise sense when $\hat{\boldsymbol{x}}_2$ is oriented toward the observer but note the sign change on the $\sin\theta_2$ relative to the other rotations)

$$R^{(2)}_{ij}(\theta_2) = \begin{bmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix} \tag{1.152}$$

and about the x_3 axis using

$$R^{(3)}_{ij}(\theta_3) = \begin{bmatrix} \cos\theta_3 & \sin\theta_3 & 0 \\ -\sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.153}$$

Now, any conceivable rotation $(\theta_1, \theta_2, \theta_3)$ is accomplished using the rotation matrix

$$R_{ij}(\theta_1, \theta_2, \theta_3) = R^{(1)}_{ik}(\theta_1)\, R^{(2)}_{kl}(\theta_2)\, R^{(3)}_{lj}(\theta_3), \tag{1.154}$$

that is, just matrix multiply the three rotation matrices together, which you can do as an end-of-chapter exercise. Tensors of any order again have Cartesian components that transform according to the above rules using $R_{ij}(\theta_1, \theta_2, \theta_3)$ as the rotation matrix, using one rotation matrix for each base vector.

As an example, the second-order identity-tensor coefficients δ_{ij} transform as $R_{ik}(\theta_1, \theta_2, \theta_3) R_{jl}(\theta_1, \theta_2, \theta_3)\delta_{kl} = R_{ik}(\theta_1, \theta_2, \theta_3)\left[R_{jk}(\theta_1, \theta_2, \theta_3)\right]^T = \delta_{ij}$, which you can confirm through direct matrix multiplication as an end-of-chapter exercise. This means that the second-order identity tensor is *isotropic*, which means that $\delta'_{ij} = \delta_{ij}$ for arbitrary coordinate rotations.

It can be convenient for proving transformation identities involving higher-order tensors if we consider small rotations $\delta\theta_i$ around each axis so that $\cos\delta\theta_i = 1 + O(\delta\theta_i^2)$ and $\sin\delta\theta_i = \delta\theta_i\left[1 + O(\delta\theta_i^2)\right]$. Upon ignoring the $O(\delta\theta_i^2)$ terms in what follows, the rotation matrix becomes

$$R_{ij}(\delta\theta_1, \delta\theta_2, \delta\theta_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \delta\theta_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + \delta\theta_2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \delta\theta_3 \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.155}$$

This can also be written as

$$R_{ij} = \delta_{ij} + \delta\theta_m \epsilon_{mij}, \tag{1.156}$$

where the ϵ_{mij} are the Levi–Civita coefficients. So, for example, if we want to find the isotropic second-order tensor that, by definition, has coefficients that satisfy $T'_{ij} = T_{ij}$, we write

$$T'_{ij} = T_{ij} = R_{ik}R_{jl}T_{kl}, \tag{1.157}$$

$$= (\delta_{ik} + \delta\theta_m \epsilon_{mik}) \left(\delta_{jl} + \delta\theta_m \epsilon_{mjl}\right) T_{kl}, \tag{1.158}$$

$$= \left[\delta_{ik}\delta_{jl} + \delta\theta_m \left(\epsilon_{mik}\delta_{jl} + \epsilon_{mjl}\delta_{ik}\right)\right] T_{kl}, \tag{1.159}$$

$$= T_{ij} + \delta\theta_m \left(\epsilon_{mik}T_{kj} + \epsilon_{mjk}T_{ik}\right). \tag{1.160}$$

This equation is satisfied if

$$\epsilon_{mik}T_{kj} = -\epsilon_{mjk}T_{ik}, \tag{1.161}$$

which has the solution $T_{ij} = T\delta_{ij}$, where T is any scalar, as can be shown through direct substitution. Thus we have that $\boldsymbol{T} = T\delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j = T\boldsymbol{I}$ is the form of the one and only isotropic second-order tensor.

If we want to find the isotropic third-order tensor that, by definition, has coefficients that satisfy $T'_{ijk} = T_{ijk}$, we write

$$T'_{ijk} = T_{ijk} = R_{il}R_{jm}R_{kn}T_{lmn}, \tag{1.162}$$

$$= \left(\delta_{il} + \delta\theta_p \epsilon_{pil}\right) \left(\delta_{jm} + \delta\theta_p \epsilon_{pjm}\right) \left(\delta_{kn} + \delta\theta_p \epsilon_{pkn}\right) T_{lmn}, \tag{1.163}$$

$$= \left[\delta_{il}\delta_{jm}\delta_{kn} + \delta\theta_p \left(\epsilon_{pil}\delta_{jm}\delta_{kn} + \epsilon_{pjm}\delta_{il}\delta_{kn} + \epsilon_{pkn}\delta_{il}\delta_{jm}\right)\right] T_{lmn} \tag{1.164}$$

$$= T_{ijk} + \delta\theta_p \left(\epsilon_{pil}T_{ljk} + \epsilon_{pjm}T_{imk} + \epsilon_{pkn}T_{ijn}\right), \tag{1.165}$$

where terms of $O(\delta\theta_i^2)$ are again ignored. This equation is satisfied if

$$\epsilon_{pil}T_{ljk} + \epsilon_{pjm}T_{imk} + \epsilon_{pkn}T_{ijn} = 0. \tag{1.166}$$

The solution of this equation is $T_{ijk} = T\epsilon_{ijk}$, where again T is any scalar, as can be seen through substitution

$$\epsilon_{pil}\epsilon_{ljk} + \epsilon_{pjl}\epsilon_{ilk} + \epsilon_{pkl}\epsilon_{ijl} = 0. \tag{1.167}$$

Rewrite this as

$$\epsilon_{lpi}\epsilon_{ljk} - \epsilon_{lpj}\epsilon_{lik} + \epsilon_{lpk}\epsilon_{lij} = 0. \tag{1.168}$$

The identity of Eq. (1.54) can then be used to write each term on the left-hand side as

$$\epsilon_{lpi}\epsilon_{ljk} = \delta_{pj}\delta_{ik} - \delta_{pk}\delta_{ij}, \tag{1.169}$$

$$-\epsilon_{lpj}\epsilon_{lik} = -\delta_{pi}\delta_{jk} + \delta_{pk}\delta_{ij}, \tag{1.170}$$

$$\epsilon_{lpk}\epsilon_{lji} = -\delta_{pj}\delta_{ik} + \delta_{pi}\delta_{jk}, \tag{1.171}$$

which sum to zero when substituted into Eq. (1.168). Thus, we have shown that ${}_3\boldsymbol{T} = T\epsilon_{ijk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k = (T)\,({}_3\boldsymbol{\epsilon})$ is the form of the one and only third-order isotropic tensor with T some arbitrary scalar.

1.8.4 Isotropic Tensors of Any Order

As just seen, isotropic tensors are those tensors whose components do not change when we change the orientation of the Cartesian coordinates. Specifically, the second-order coefficients δ_{ij} and third-order coefficients ϵ_{ijk} do not change when changing the orientation of the coordinates. As such, even-ordered isotropic tensors have coefficients that are multiples of the Kronecker coefficients δ_{ij} and odd-ordered isotropic tensors have coefficients that involve the single presence of the Levi–Civita coefficients ϵ_{ijk} and additional multiples of the Kronecker coefficients that get to the desired (odd) tensorial order. So higher-order isotropic tensors involve multiples of Kronecker and Levi–Civita coefficients with numbers of indices that add up to the tensorial order (or rank) of interest.

We call a zeroth-order tensor a scalar and all scalars are, by definition, independent of the orientation of the axes. So all scalars are isotropic.

We call a first-order tensor a vector and all vectors of finite length have components that change when the axes are rotated. So an "isotropic vector" has zero length and does not exist.

As proven above, there is one fundamental *second-order isotropic tensor* ${}_2\boldsymbol{I}$, which is the second-order identity tensor $\boldsymbol{I}$,

$${}_2\boldsymbol{I} = \boldsymbol{I} = \delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j. \tag{1.172}$$

Similarly, we showed there is one fundamental *third-order isotropic tensor* ${}_3\boldsymbol{I}$, which is the third-order Levi–Civita alternating (or "antisymmetric" or "permutation") tensor,

$${}_3\boldsymbol{I} = {}_3\boldsymbol{\epsilon} = \epsilon_{ijk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k. \tag{1.173}$$

We can multiply these fundamental second-order and third-order isotropic tensors by scalars, and the result will also be an isotropic tensor.

Higher-order isotropic tensors have coefficients that involve additional multiples of the Kronecker coefficients. So, for example, there are multiple fundamental *fourth-order isotropic tensors* ${}_4\boldsymbol{I}^{(m)}$. If we define the first dummy index of these coefficients to always be i, there are three unique ways to place the remaining j, k, and l indices across the a, b, and c positions of $\delta_{ia}\delta_{bc}$, that is,

$$ {}_4\boldsymbol{I}^{(1)} = \delta_{il}\delta_{jk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l, \tag{1.174} $$

$$ {}_4\boldsymbol{I}^{(2)} = \delta_{ik}\delta_{jl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l, \tag{1.175} $$

$$ {}_4\boldsymbol{I}^{(3)} = \delta_{ij}\delta_{kl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l = \boldsymbol{II}. \tag{1.176} $$

There are not more than these three fundamental fourth-order isotropic tensors due to the symmetry $\delta_{ij} = \delta_{ji}$. The rule for the number M of fundamental even-ordered $n = 4, 6, 8, \ldots$ isotropic tensors involving only the Kronecker coefficients is

$$ M = \prod_{i=1}^{n/2} (n - 2i + 1) \quad \text{for even } n. \tag{1.177} $$

So for $n = 4$, we have $M = (n-1)(n-3) = (3)(1) = 3$ as seen in Eq. (1.177).

The fundamental *fifth-order isotropic tensors* ${}_5\boldsymbol{I}^{(m)}$ have coefficients that involve a single multiplication between the Levi–Civita coefficients and the Kronecker coefficients. If we define the first dummy index of these fifth-order coefficients as i, there are six unique nonzero ways to place the remaining j, k, l, m indices across the a, b, c, d positions of $\epsilon_{iab}\delta_{cd}$ and four unique nonzero way to place the j, k, l, m across the a, b, c, d positions of $\delta_{ia}\epsilon_{bcd}$ to give

$$ {}_5\boldsymbol{I}^{(1)} = \epsilon_{ijk}\delta_{lm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m = ({}_3\boldsymbol{\epsilon})\,\boldsymbol{I}, \tag{1.178} $$

$$ {}_5\boldsymbol{I}^{(2)} = \epsilon_{ijl}\delta_{km}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.179} $$

$$ {}_5\boldsymbol{I}^{(3)} = \epsilon_{ijm}\delta_{kl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.180} $$

$$ {}_5\boldsymbol{I}^{(4)} = \epsilon_{ikl}\delta_{jm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.181} $$

$$ {}_5\boldsymbol{I}^{(5)} = \epsilon_{ikm}\delta_{jl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.182} $$

$$ {}_5\boldsymbol{I}^{(6)} = \epsilon_{ilm}\delta_{jk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.183} $$

$$ {}_5\boldsymbol{I}^{(7)} = \delta_{ij}\epsilon_{klm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m = \boldsymbol{I}\,({}_3\boldsymbol{\epsilon}), \tag{1.184} $$

$$ {}_5\boldsymbol{I}^{(8)} = \delta_{ik}\epsilon_{jlm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.185} $$

$$ {}_5\boldsymbol{I}^{(9)} = \delta_{il}\epsilon_{jkm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.186} $$

$$ {}_5\boldsymbol{I}^{(10)} = \delta_{im}\epsilon_{jkl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m, \tag{1.187} $$

There are not more than these 10 fundamental fifth-order isotropic tensors because $\delta_{ij} = \delta_{ji}$ and $\epsilon_{ijk} = -\epsilon_{ikj} = -\epsilon_{kij}$, that is, multiplying by -1 does not create a distinct fundamental isotropic tensor.

Proceeding like above gives the 15 *sixth-order isotropic tensors* ${}_6\boldsymbol{I}^{(m)}$ involving the Kronecker coefficients:

$$
\begin{aligned}
{}_6\boldsymbol{I}^{(1)} &= \delta_{in}\delta_{ml}\delta_{jk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.188)\\
{}_6\boldsymbol{I}^{(2)} &= \delta_{in}\delta_{mk}\delta_{jl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.189)\\
{}_6\boldsymbol{I}^{(3)} &= \delta_{in}\delta_{mj}\delta_{kl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.190)\\
{}_6\boldsymbol{I}^{(4)} &= \delta_{im}\delta_{nl}\delta_{jk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.191)\\
{}_6\boldsymbol{I}^{(5)} &= \delta_{im}\delta_{nk}\delta_{jl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.192)\\
{}_6\boldsymbol{I}^{(6)} &= \delta_{im}\delta_{nj}\delta_{kl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.193)\\
{}_6\boldsymbol{I}^{(7)} &= \delta_{il}\delta_{nm}\delta_{jk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n = \left({}_4\boldsymbol{I}^{(1)}\right)\boldsymbol{I}, && (1.194)\\
{}_6\boldsymbol{I}^{(8)} &= \delta_{il}\delta_{nk}\delta_{jm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.195)\\
{}_6\boldsymbol{I}^{(9)} &= \delta_{il}\delta_{nj}\delta_{km}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.196)\\
{}_6\boldsymbol{I}^{(10)} &= \delta_{ik}\delta_{nm}\delta_{jl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n = \left({}_4\boldsymbol{I}^{(2)}\right)\boldsymbol{I} && (1.197)\\
{}_6\boldsymbol{I}^{(11)} &= \delta_{ik}\delta_{nl}\delta_{jm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.198)\\
{}_6\boldsymbol{I}^{(12)} &= \delta_{ik}\delta_{nj}\delta_{lm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n, && (1.199)\\
{}_6\boldsymbol{I}^{(13)} &= \delta_{ij}\delta_{nm}\delta_{kl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n = \left({}_4\boldsymbol{I}^{(3)}\right)\boldsymbol{I} = \boldsymbol{I}\left({}_4\boldsymbol{I}^{(3)}\right) = \boldsymbol{III}, && (1.200)\\
{}_6\boldsymbol{I}^{(14)} &= \delta_{ij}\delta_{nl}\delta_{km}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n = \boldsymbol{I}\left({}_4\boldsymbol{I}^{(2)}\right), && (1.201)\\
{}_6\boldsymbol{I}^{(15)} &= \delta_{ij}\delta_{nk}\delta_{lm}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n = \boldsymbol{I}\left({}_4\boldsymbol{I}^{(1)}\right). && (1.202)
\end{aligned}
$$

Using the rule of Eq. (1.177) for this case of order $n = 6$, we have $M = (n-1)(n-3)(n-5) = (5)(3)(1) = 15$ as the number of fundamental sixth-order isotropic tensors involving only the Kronecker coefficients. To these can be added the sixteenth and final sixth-order isotropic tensor

$$
{}_6\boldsymbol{I}^{(16)} = \epsilon_{ijk}\epsilon_{lmn}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n = ({}_3\boldsymbol{\epsilon})\,({}_3\boldsymbol{\epsilon}) \qquad (1.203)
$$

with the ϵ_{ijk} the Levi–Civita coefficients.

If you want to determine the seventh-order isotropic tensors, arrange the j, k, l, m, n, o indices in the a, b, c, d, e, f positions of the following coefficients: $\epsilon_{iab}\delta_{cd}\delta_{ef}$, then $\delta_{ia}\epsilon_{bcd}\delta_{ef}$ and finally $\delta_{ia}\delta_{bc}\epsilon_{def}$. However, we will not carry out this exercise because the highest-order isotropic tensor we will encounter in our treatment of constitutive laws in this book is the sixth order as given in Section 4.1.5.

1.8.5 Curie's Principle for the Constitutive Laws of Isotropic Media

Curie's principle (Curie, 1894), as named after physicist Pierre Curie, has been a source of controversy over the years but will be taken here to be the noncontroversial statement (theorem, in fact, as will be demonstrated below) that says: "*in a constitutive law of an isotropic material, a generalized response has the same tensorial order as the generalized forces that create it*." Curie's (1894) main point is actually the corollary statement that if a response and a force in a constitutive law are to have different tensorial orders, the material must possess *anisotropy*, that is, cannot be purely isotropic. Although we are getting ahead of ourselves in terms of the physics, we will clarify the meaning of these various words and prove the above italicized statement now, rather than in later applications, because the demonstration comes directly from the nature of the isotropic tensors that were just determined in Sections 1.8.3 and 1.8.4.

We focus here on constitutive laws associated with reversible processes, but Curie's principle also applies to irreversible processes for which the constitutive laws are called transport laws (Chapter 7). If a "force" (or "cause") is applied to an element of matter to create some "response" (or "effect"), the process is called *reversible* if when the force is returned to its initial value, the response returns to its initial value. Elastic deformation and electric and magnetic polarization are examples of reversible processes as will be developed from first principles in Chapters 3 and 4. For reversible processes, the constitutive laws are always temporal differential equations and, as developed in Chapter 6 on the thermodynamics of reversible processes, are derived by taking total time derivatives of a scalar "fundamental function" that we define here in generic form to be $u = u(\alpha, \boldsymbol{a}, \boldsymbol{A})$. This scalar function depends on time-variable forces that we represent here as a scalar $\alpha(t)$, a vector $\boldsymbol{a}(t)$, and a second-order tensor $\boldsymbol{A}(t)$ that is always symmetric. When we develop the physical nature of such a fundamental function in the chapters that follow, the function u will be seen to represent the internal energy of an element (defined later), while the scalar α can be representing entropy (defined later), the vector $\boldsymbol{a}$ can be representing the dielectric displacement or applied electric field (defined later) and the second-order tensor $\boldsymbol{A}$ is representing the elastic deformation or strain tensor (defined later) with $\boldsymbol{A}$ being symmetric. However, such physical interpretations are not required in our proof of Curie's principle that only requires a function $u = u(\alpha, \boldsymbol{a}, \boldsymbol{A})$ with $\boldsymbol{A}$ symmetric and knowledge about isotropic tensors.

Begin the proof by taking a total time derivative of the given fundamental function $u = u(\alpha, \boldsymbol{a}, \boldsymbol{A})$ to obtain

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \left(\frac{\partial u}{\partial \alpha}\right)\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \left(\frac{\partial u}{\partial \boldsymbol{a}}\right)\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + \left(\frac{\partial u}{\partial \boldsymbol{A}}\right):\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t}. \tag{1.204}$$

The partial derivatives in brackets are called the "responses" to which we give the symbolic names

$$\beta = \frac{\partial u}{\partial \alpha} \tag{1.205}$$

$$\boldsymbol{b} = \frac{\partial u}{\partial \boldsymbol{a}} \mathrel{\widehat{=}} \frac{\partial u}{\partial a_i}\hat{\boldsymbol{x}}_i, \tag{1.206}$$

$$\boldsymbol{B} = \frac{\partial u}{\partial \boldsymbol{A}} \mathrel{\widehat{=}} \frac{\partial u}{\partial A_{ij}}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j. \tag{1.207}$$

The second statements for both $\boldsymbol{b}$ and $\boldsymbol{B}$ define what it means to take a partial derivative when the independent variable is a vector or tensor. In later development, we will see that β is representing temperature if α is entropy, $\boldsymbol{b}$ is the total electric field that includes polarization if $\boldsymbol{a}$ is the applied electric field (dielectric displacement), $\boldsymbol{B}$ is the stress tensor if $\boldsymbol{A}$ is the strain tensor and Eq. (1.204) is the first law of thermodynamics. But again, such physical identifications are not required in this proof of Curie's principle, that simply posits the existence of a fundamental function $u = u(\alpha, \boldsymbol{a}, \boldsymbol{A})$ with $\boldsymbol{A}$ being symmetric.

The "constitutive laws" are the total time derivatives of the generalized responses β, $\boldsymbol{b}$, and $\boldsymbol{B}$:

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = \left(\frac{\partial^2 u}{\partial\alpha^2}\right)\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \left(\frac{\partial^2 u}{\partial\boldsymbol{a}\partial\alpha}\right)\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + \left(\frac{\partial^2 u}{\partial\boldsymbol{A}\partial\alpha}\right):\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}, \tag{1.208}$$

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t} = \left(\frac{\partial^2 u}{\partial\alpha\partial\boldsymbol{a}}\right)\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \left(\frac{\partial^2 u}{\partial\boldsymbol{a}\partial\boldsymbol{a}}\right)\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + \left(\frac{\partial^2 u}{\partial\boldsymbol{A}\partial\boldsymbol{a}}\right):\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}, \tag{1.209}$$

$$\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t} = \left(\frac{\partial^2 u}{\partial\alpha\partial\boldsymbol{A}}\right)\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \left(\frac{\partial^2 u}{\partial\boldsymbol{a}\partial\boldsymbol{A}}\right)\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + \left(\frac{\partial^2 u}{\partial\boldsymbol{A}\partial\boldsymbol{A}}\right):\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}, \tag{1.210}$$

where the various double derivatives of the fundamental function having different tensorial orders are called "material properties" and can be given the symbolic names that possess the following symmetries:

$$\gamma = \frac{\partial\beta}{\partial\alpha} = \frac{\partial^2 u}{\partial\alpha^2}, \tag{1.211}$$

$$\boldsymbol{c} = \frac{\partial\beta}{\partial\boldsymbol{a}} = \frac{\partial^2 u}{\partial\boldsymbol{a}\partial\alpha} = \frac{\partial^2 u}{\partial\alpha\partial\boldsymbol{a}} = \frac{\partial\boldsymbol{b}}{\partial\alpha}, \tag{1.212}$$

$$\boldsymbol{D} = \frac{\partial\beta}{\partial\boldsymbol{A}} = \frac{\partial^2 u}{\partial\boldsymbol{A}\partial\alpha} = \frac{\partial^2 u}{\partial\alpha\partial\boldsymbol{A}} = \frac{\partial\boldsymbol{B}}{\partial\alpha} = \boldsymbol{D}^T, \tag{1.213}$$

$$\boldsymbol{E} = \frac{\partial\boldsymbol{b}}{\partial\boldsymbol{a}} = \frac{\partial^2 u}{\partial\boldsymbol{a}\partial\boldsymbol{a}} = \boldsymbol{E}^T, \tag{1.214}$$

$${}_3\boldsymbol{F} = \frac{\partial\boldsymbol{b}}{\partial\boldsymbol{A}} = \frac{\partial^2 u}{\partial\boldsymbol{A}\partial\boldsymbol{a}} = \left(\frac{\partial^2 u}{\partial\boldsymbol{a}\partial\boldsymbol{A}}\right)^{\overset{T}{231}} = {}_3\boldsymbol{F}^{\overset{T}{213}}, \tag{1.215}$$

$${}_3\boldsymbol{F}^{\overset{T}{312}} = \frac{\partial\boldsymbol{B}}{\partial\boldsymbol{a}} = \frac{\partial^2 u}{\partial\boldsymbol{a}\partial\boldsymbol{A}} = \left(\frac{\partial^2 u}{\partial\boldsymbol{A}\partial\boldsymbol{a}}\right)^{\overset{T}{312}} = {}_3\boldsymbol{F}^{\overset{T}{321}}, \tag{1.216}$$

$${}_4\boldsymbol{G} = \frac{\partial\boldsymbol{B}}{\partial\boldsymbol{A}} = \frac{\partial^2 u}{\partial\boldsymbol{A}\partial\boldsymbol{A}} = {}_4\boldsymbol{G}^{\overset{T}{3412}} = {}_4\boldsymbol{G}^{\overset{T}{2134}} = {}_4\boldsymbol{G}^{\overset{T}{1243}}. \tag{1.217}$$

So in terms of the material properties γ (a scalar), $\boldsymbol{c}$ (a vector), $\boldsymbol{D}$ (a second-order tensor), $\boldsymbol{E}$ (a second-order tensor), ${}_3\boldsymbol{F}$ (a third-order tensor), and ${}_4\boldsymbol{G}$ (a fourth-order tensor), the constitutive laws can be written

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = \gamma\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \boldsymbol{c}\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + \boldsymbol{D}:\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}, \tag{1.218}$$

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t} = \boldsymbol{c}\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \boldsymbol{E}\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + {}_3\boldsymbol{F}:\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}, \tag{1.219}$$

$$\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t} = \boldsymbol{D}\frac{\mathrm{d}\alpha}{\mathrm{d}t} + {}_3\boldsymbol{F}^{\overset{T}{312}}\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t} + {}_4\boldsymbol{G}:\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}. \tag{1.220}$$

Because such total differentials can be integrated reversibly, these laws correspond to "reversible processes." So a response of a given tensorial order can, in general, be generated by a force of different tensorial order.

For the material to be called "isotropic," the coefficients of each tensorial material property must be invariant to rotations of the coordinates, that is, each material property must involve a scalar times the fundamental isotropic tensor(s) having the same tensorial order as the material property. So in an isotropic material, we must have

$$\boldsymbol{c} = 0 \text{ (because there are no isotropic vectors)}, \tag{1.221}$$

$$\boldsymbol{D} = d\,\delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j = d\boldsymbol{I} \text{ (with } d \text{ a scalar)}, \tag{1.222}$$

$$\boldsymbol{E} = e\,\delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j = e\boldsymbol{I} \text{ (with } e \text{ a scalar)}, \tag{1.223}$$

$${}_3\boldsymbol{F} = f\epsilon_{ijk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k = (f)\,({}_3\boldsymbol{\epsilon}) \text{ (with } f \text{ a scalar)}, \tag{1.224}$$

$${}_3\boldsymbol{F}^{\overset{T}{312}} = {}_3\boldsymbol{F} \text{ (because } \epsilon_{kij} = \epsilon_{ijk}), \tag{1.225}$$

$${}_4\boldsymbol{G} = g_1\left({}_4\boldsymbol{I}^{(1)} + {}_4\boldsymbol{I}^{(2)}\right) + (g_2)\,{}_4\boldsymbol{I}^{(3)} \text{ (with } g_1 \text{ and } g_2 \text{ scalars)}. \tag{1.226}$$

The three fundamental fourth-order isotropic tensors ${}_4\boldsymbol{I}^{(1)}$, ${}_4\boldsymbol{I}^{(2)}$, and ${}_4\boldsymbol{I}^{(3)}$ are given in Eqs (1.174)–(1.176). As shown in an end-of-chapter exercise, we have ${}_4\boldsymbol{I}^{(1)} : \mathrm{d}\boldsymbol{A}/\mathrm{d}t = \mathrm{d}\boldsymbol{A}/\mathrm{d}t$, ${}_4\boldsymbol{I}^{(2)} : \mathrm{d}\boldsymbol{A}/\mathrm{d}t = \mathrm{d}\boldsymbol{A}^T/\mathrm{d}t = \mathrm{d}\boldsymbol{A}/\mathrm{d}t$, and ${}_4\boldsymbol{I}^{(3)} : \mathrm{d}\boldsymbol{A}/\mathrm{d}t = \left[\mathrm{d}(\boldsymbol{I}:\boldsymbol{A})/\mathrm{d}t\right]\boldsymbol{I}$, where $\boldsymbol{I}$ is the second-order identity tensor.

Because ${}_3\boldsymbol{F}$ is proportional to the Levi–Civita tensor, when it is double dotted into the symmetric tensor $\mathrm{d}\boldsymbol{A}/\mathrm{d}t$ we get zero as proven earlier. When this same ${}_3\boldsymbol{F}$ is dotted into the vector $\mathrm{d}\boldsymbol{a}/\mathrm{d}t$, we obtain an antisymmetric tensor, also as proven earlier. However, $\mathrm{d}\boldsymbol{B}/\mathrm{d}t$ is a symmetric tensor because $\boldsymbol{A}$ is symmetric, which tells us that the scalar material property f must be zero. Last, we can decompose the second-order tensor $\boldsymbol{A}$ into so-called *isotropic* and *deviatoric* portions as

$$\boldsymbol{A} = \underbrace{\left(\frac{\boldsymbol{I}:\boldsymbol{A}}{\boldsymbol{I}:\boldsymbol{I}}\right)\boldsymbol{I}}_{\textit{Isotropic portion}} + \underbrace{\boldsymbol{A} - \left(\frac{\boldsymbol{I}:\boldsymbol{A}}{\boldsymbol{I}:\boldsymbol{I}}\right)\boldsymbol{I}}_{\textit{Deviatoric portion } \boldsymbol{A}^D}, \tag{1.227}$$

and similarly for $\boldsymbol{B} = (\boldsymbol{I}:\boldsymbol{B}/\boldsymbol{I}:\boldsymbol{I})\,\boldsymbol{I} + \boldsymbol{B}^D$. The deviatoric or "true-tensorial" portion of a second-order tensor has zero trace, that is, $\boldsymbol{I}:\boldsymbol{A}^D = 0 = \boldsymbol{I}:\boldsymbol{B}^D$ and continues to be symmetric if the second-order tensor being decomposed is symmetric. Note that if $\boldsymbol{B}$ is the stress tensor, as defined in Chapter 2, then $-(\boldsymbol{I}:\boldsymbol{B})/(\boldsymbol{I}:\boldsymbol{I})$ is the scalar pressure.

Thus, in an isotropic material, the constitutive laws contained within $u = u(\alpha, \boldsymbol{a}, \boldsymbol{A})$ are

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = \gamma\frac{\mathrm{d}\alpha}{\mathrm{d}t} + d\,\frac{\mathrm{d}(\boldsymbol{I}:\boldsymbol{A})}{\mathrm{d}t}, \tag{1.228}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\boldsymbol{I}:\boldsymbol{B}}{\boldsymbol{I}:\boldsymbol{I}}\right) = d\,\frac{\mathrm{d}\alpha}{\mathrm{d}t} + \left(g_2 + \frac{2g_1}{\boldsymbol{I}:\boldsymbol{I}}\right)\frac{\mathrm{d}(\boldsymbol{I}:\boldsymbol{A})}{\mathrm{d}t}, \tag{1.229}$$

$$\frac{\mathrm{d}\boldsymbol{b}}{\mathrm{d}t} = e\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}t}, \tag{1.230}$$

$$\frac{\mathrm{d}\boldsymbol{B}^D}{\mathrm{d}t} = g_1\frac{\mathrm{d}\boldsymbol{A}^D}{\mathrm{d}t}. \tag{1.231}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

We see that the time rate of each "response" in these isotropic laws (the left-hand side) has the same tensorial order as the time rate of the conjugate "forces" that are creating it (the right-hand side) and that all the material properties are now simple scalars (γ, d, e, g_1, and g_2 with $f = 0$). This is the content of *Curie's principle of constitutive laws in isotropic media* that we have now demonstrated to be a theorem for the constitutive laws contained in the fundamental function $u = u(\alpha, \boldsymbol{a}, \boldsymbol{A})$ with $\boldsymbol{A}$ symmetric. So, for example, if a vectorial response $\boldsymbol{b}$ is created by a second-order tensor $\boldsymbol{A}$ as controlled by the third-order material property ${}_3\boldsymbol{F}$ (piezoelectricity is a classic example), the material cannot be isotropic, that is, it must possess *anisotropy* so that ${}_3\boldsymbol{F}$ is not isotropic, which is the main message Curie (1894) was conveying.

1.8.6 Tensor Calculus in Orthogonal Curvilinear Coordinates

We now present the detailed expressions for various common tensor-calculus operations involving the gradient operator in orthogonal curvilinear coordinates. The various tensor-calculus operations are also given explicitly in both cylindrical and spherical coordinates, which are the two most commonly employed curvilinear coordinates you will encounter and the only curvilinear coordinates used in this book. The treatment that follows is inspired by the fabulous treatment of orthogonal curvilinear coordinates in Appendix A of the fluid-mechanics text by Happel and Brenner (1983).

To begin, consider the differences between Cartesian coordinates and some arbitrary orthogonal curvilinear coordinate system as shown in Fig. 1.14. The distance vector in Cartesian coordinates is written in array format as $\boldsymbol{r} = (x_1, x_2, x_3)$. Because each coordinate in Cartesians measures linear distance along that coordinate, we have that the infinitesimal distance vector $\mathrm{d}\boldsymbol{r}$ between two positions in space $\mathrm{d}\boldsymbol{r} = \hat{\boldsymbol{x}}_1\, \mathrm{d}\ell_1 + \hat{\boldsymbol{x}}_2\, \mathrm{d}\ell_2 + \hat{\boldsymbol{x}}_3\, \mathrm{d}\ell_3$ is written

$$\mathrm{d}\boldsymbol{r} = \hat{\boldsymbol{x}}_1\, \mathrm{d}x_1 + \hat{\boldsymbol{x}}_2\, \mathrm{d}x_2 + \hat{\boldsymbol{x}}_3\, \mathrm{d}x_3 \tag{1.232}$$

because $\mathrm{d}\ell_i = \mathrm{d}x_i$ in each direction i. Thus, the gradient operator is simply

$$\nabla = \hat{\boldsymbol{x}}_1 \frac{\partial}{\partial x_1} + \hat{\boldsymbol{x}}_2 \frac{\partial}{\partial x_2} + \hat{\boldsymbol{x}}_3 \frac{\partial}{\partial x_3}, \tag{1.233}$$

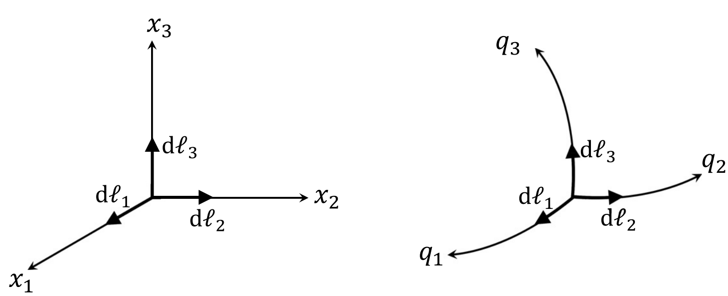

Figure 1.14 Cartesian coordinates on the left and some orthogonal curvilinear coordinate system on the right.

where we can employ the notation, if we so choose, that $\nabla = \partial/\partial \boldsymbol{r}$ with the Eq. (1.233) interpretation of what $\partial/\partial \boldsymbol{r}$ means to do.

The situation is different in orthogonal curvilinear coordinates because the coordinates in each orthogonal direction do not necessarily represent distance but instead can be represented by angles and because the base vectors vary their orientation, in general, as we move along a given coordinate. Although some point in space can again be represented in array format as $\boldsymbol{r} = (q_1, q_2, q_3)$, the infinitesimal distance between two points in space $\mathrm{d}\boldsymbol{r} = \hat{\boldsymbol{q}}_1\,\mathrm{d}\ell_1 + \hat{\boldsymbol{q}}_2\,\mathrm{d}\ell_2 + \hat{\boldsymbol{q}}_3\,\mathrm{d}\ell_3$ is now

$$\mathrm{d}\boldsymbol{r} = \hat{\boldsymbol{q}}_1 \frac{\mathrm{d}q_1}{h_1} + \hat{\boldsymbol{q}}_2 \frac{\mathrm{d}q_2}{h_2} + \hat{\boldsymbol{q}}_3 \frac{\mathrm{d}q_3}{h_3}, \tag{1.234}$$

that is, infinitesimal distance in each coordinate direction is given by

$$\mathrm{d}\ell_i = \frac{\mathrm{d}q_i}{h_i(q_1, q_2, q_3)}, \tag{1.235}$$

where the coefficients h_i are called the *metrical coefficients* for the particular orthogonal curvilinear coordinate system under consideration. These metrical coefficients convert change along a coordinate direction to change in distance along that coordinate and themselves will vary through space in general. *Specifying a curvilinear coordinate system amounts to specifying the functional dependence of the three metrical coefficients on the coordinates q_1, q_2, and q_3.*

The gradient operator in orthogonal curvilinear coordinates is then

$$\nabla = \hat{\boldsymbol{q}}_1\, h_1 \frac{\partial}{\partial q_1} + \hat{\boldsymbol{q}}_2\, h_2 \frac{\partial}{\partial q_2} + \hat{\boldsymbol{q}}_3\, h_3 \frac{\partial}{\partial q_3}. \tag{1.236}$$

A useful definition for the base vectors comes from combining $\mathrm{d}\boldsymbol{r} = \hat{\boldsymbol{q}}_1\,\mathrm{d}\ell_1 + \hat{\boldsymbol{q}}_2\,\mathrm{d}\ell_2 + \hat{\boldsymbol{q}}_3\,\mathrm{d}\ell_3$ with Eq. (1.235)

$$\hat{\boldsymbol{q}}_i = h_i \frac{\partial \boldsymbol{r}}{\partial q_i} \quad \text{where } i = 1, 2, \text{ or } 3 \tag{1.237}$$

and where there is no summation here over the repeated index on the right-hand side. Note that $\partial q_i/\partial q_j = 1$ when $i = j$ but $\partial q_i/\partial q_j = 0$ when $i \neq j$. Cartesian coordinates are defined by taking $h_i = 1$.

With this introduction to the metrical coefficients $h_i(q_1, q_2, q_3)$, we next use that the coordinates (q_1, q_2, q_3), though curvilinear, are also orthogonal to each other at each point. This means $\hat{\boldsymbol{q}}_i \cdot \hat{\boldsymbol{q}}_j = 0$ when $i \neq j$. Similarly, $\hat{\boldsymbol{q}}_i = \hat{\boldsymbol{q}}_j \times \hat{\boldsymbol{q}}_k$ where the indices are ordered here in the right-handed sense of $[ijk] = [123]$, $[231]$, or $[312]$. Introducing Eq. (1.237) into these statements of orthogonality and taking the partial derivative with respect to each coordinate q_i, one arrives, eventually, at the following results for the derivatives of the base vectors in an orthogonal curvilinear coordinate system

$$\frac{\partial \hat{\boldsymbol{q}}_j}{\partial q_i} = \hat{\boldsymbol{q}}_i h_j \frac{\partial}{\partial q_j}\left(\frac{1}{h_i}\right) \qquad \text{where } j \neq i \tag{1.238}$$

$$\frac{\partial \hat{\boldsymbol{q}}_i}{\partial q_i} = -\hat{\boldsymbol{q}}_j h_j \frac{\partial}{\partial q_j}\left(\frac{1}{h_i}\right) - \hat{\boldsymbol{q}}_k h_k \frac{\partial}{\partial q_k}\left(\frac{1}{h_i}\right) \qquad \text{where } j \neq k \neq i \tag{1.239}$$

and where there is no summation over repeated indices in these expressions. Again, in Cartesian coordinates, all of these derivatives are zero.

The second-order identity tensor $\boldsymbol{I}$ in orthogonal curvilinear coordinates is defined

$$\boldsymbol{I} = \nabla \boldsymbol{r} = \hat{\boldsymbol{q}}_1\, h_1 \frac{\partial \boldsymbol{r}}{\partial q_1} + \hat{\boldsymbol{q}}_2\, h_2 \frac{\partial \boldsymbol{r}}{\partial q_2} + \hat{\boldsymbol{q}}_3\, h_3 \frac{\partial \boldsymbol{r}}{\partial q_3}. \tag{1.240}$$

which from Eq. (1.237) is simply

$$\boldsymbol{I} = \hat{\boldsymbol{q}}_1\hat{\boldsymbol{q}}_1 + \hat{\boldsymbol{q}}_2\hat{\boldsymbol{q}}_2 + \hat{\boldsymbol{q}}_3\hat{\boldsymbol{q}}_3 = \delta_{ij}\hat{\boldsymbol{q}}_i\hat{\boldsymbol{q}}_j, \tag{1.241}$$

just like in Cartesian coordinates.

We also have that infinitesimal surface elements $\mathrm{d}S_i$ having a normal in the $\hat{\boldsymbol{q}}_i$ direction are given by

$$\mathrm{d}S_1 = \mathrm{d}\ell_2 \mathrm{d}\ell_3 = \frac{\mathrm{d}q_2 \mathrm{d}q_3}{h_2 h_3} \tag{1.242}$$

$$\mathrm{d}S_2 = \mathrm{d}\ell_1 \mathrm{d}\ell_3 = \frac{\mathrm{d}q_1 \mathrm{d}q_3}{h_1 h_3} \tag{1.243}$$

$$\mathrm{d}S_3 = \mathrm{d}\ell_1 \mathrm{d}\ell_2 = \frac{\mathrm{d}q_1 \mathrm{d}q_2}{h_1 h_2}. \tag{1.244}$$

Similarly, the infinitesimal volume element is

$$\mathrm{d}V = \mathrm{d}\ell_1 \mathrm{d}\ell_2 \mathrm{d}\ell_3 = \frac{\mathrm{d}q_1 \mathrm{d}q_2 \mathrm{d}q_3}{h_1 h_2 h_3}. \tag{1.245}$$

So given the metrical coefficients for an orthogonal curvilinear coordinate system, we can now calculate spatial derivatives of vectors and set up surface and volume integrals in those coordinates.

For example, to perform the divergence of a vector field $\nabla \cdot \boldsymbol{a}$, we write

$$\nabla \cdot \boldsymbol{a} = \hat{\boldsymbol{q}}_i\, h_i \frac{\partial}{\partial q_i} \cdot (\hat{\boldsymbol{q}}_j a_j) \tag{1.246}$$

$$= h_i \hat{\boldsymbol{q}}_i \cdot \left[\left(\frac{\partial \hat{\boldsymbol{q}}_j}{\partial q_i}\right) a_j + \hat{\boldsymbol{q}}_j \frac{\partial a_j}{\partial q_i}\right] \tag{1.247}$$

where now there is summation assumed over repeated indices and, as earlier, we use a different index for each vector in the tensor-calculus expression to be evaluated prior to performing any scalar products. We perform the explicit sum over repeated indices, insert Eqs (1.238) and (1.239) for the various derivatives of the base vectors and use the orthogonality condition that $\hat{\boldsymbol{q}}_i \cdot \hat{\boldsymbol{q}}_j = \delta_{ij}$, which is nonzero only if $j = i$, to obtain

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\nabla \cdot \boldsymbol{a} = h_1 h_2 h_3 \left[\frac{\partial}{\partial q_1} \left(\frac{a_1}{h_2 h_3} \right) + \frac{\partial}{\partial q_2} \left(\frac{a_2}{h_1 h_3} \right) + \frac{\partial}{\partial q_3} \left(\frac{a_3}{h_1 h_2} \right) \right]. \tag{1.248}$$

This then yields the Laplacian $\nabla^2 \psi \widehat{=} \nabla \cdot \nabla \psi$ of any scalar field ψ to be

$$\nabla^2 \psi = h_1 h_2 h_3 \times \left[\frac{\partial}{\partial q_1} \left(\frac{h_1 \partial \psi / \partial q_1}{h_2 h_3} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_2 \partial \psi / \partial q_2}{h_1 h_3} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_3 \partial \psi / \partial q_3}{h_1 h_2} \right) \right]. \tag{1.249}$$

Similar operations yield the curl in the form

$$\begin{aligned} \nabla \times \boldsymbol{a} = {} & \hat{\boldsymbol{q}}_1 h_2 h_3 \left[\frac{\partial}{\partial q_2} \left(\frac{a_3}{h_3} \right) - \frac{\partial}{\partial q_3} \left(\frac{a_2}{h_2} \right) \right] \\ & + \hat{\boldsymbol{q}}_2 h_1 h_3 \left[\frac{\partial}{\partial q_3} \left(\frac{a_1}{h_1} \right) - \frac{\partial}{\partial q_1} \left(\frac{a_3}{h_3} \right) \right] \\ & + \hat{\boldsymbol{q}}_3 h_1 h_2 \left[\frac{\partial}{\partial q_1} \left(\frac{a_2}{h_2} \right) - \frac{\partial}{\partial q_2} \left(\frac{a_1}{h_1} \right) \right]. \end{aligned} \tag{1.250}$$

The second-order tensor $\nabla \boldsymbol{a}$ in orthogonal curvilinear coordinates is defined (again with summation over repeated indices)

$$\nabla \boldsymbol{a} = \hat{\boldsymbol{q}}_i \, h_i \frac{\partial}{\partial q_i} \left(\hat{\boldsymbol{q}}_j a_j \right) = \hat{\boldsymbol{q}}_i \, h_i \left[\left(\frac{\partial \hat{\boldsymbol{q}}_j}{\partial q_i} \right) a_j + \hat{\boldsymbol{q}}_j \frac{\partial a_j}{\partial a_i} \right]. \tag{1.251}$$

So performing the explicit sum over repeated indices and employing Eqs (1.238) and (1.239) for the derivatives of the base vectors, we obtain the nine components of $\nabla \boldsymbol{a}$ as

$$\begin{aligned} \nabla \boldsymbol{a} = {} & \hat{\boldsymbol{q}}_1 \hat{\boldsymbol{q}}_1 h_1 \left[\frac{\partial a_1}{\partial q_1} + h_2 a_2 \frac{\partial}{\partial q_2} \left(\frac{1}{h_1} \right) + h_3 a_3 \frac{\partial}{\partial q_3} \left(\frac{1}{h_1} \right) \right] \\ & + \hat{\boldsymbol{q}}_1 \hat{\boldsymbol{q}}_2 h_1 \left[\frac{\partial a_2}{\partial q_1} - h_2 a_1 \frac{\partial}{\partial q_2} \left(\frac{1}{h_1} \right) \right] \\ & + \hat{\boldsymbol{q}}_1 \hat{\boldsymbol{q}}_3 h_1 \left[\frac{\partial a_3}{\partial q_1} - h_3 a_1 \frac{\partial}{\partial q_3} \left(\frac{1}{h_1} \right) \right] \\ & + \hat{\boldsymbol{q}}_2 \hat{\boldsymbol{q}}_1 h_2 \left[\frac{\partial a_1}{\partial q_2} - h_1 a_2 \frac{\partial}{\partial q_1} \left(\frac{1}{h_2} \right) \right] \\ & + \hat{\boldsymbol{q}}_2 \hat{\boldsymbol{q}}_2 h_2 \left[\frac{\partial a_2}{\partial q_2} + h_3 a_3 \frac{\partial}{\partial q_3} \left(\frac{1}{h_2} \right) + h_1 a_1 \frac{\partial}{\partial q_1} \left(\frac{1}{h_2} \right) \right] \\ & + \hat{\boldsymbol{q}}_2 \hat{\boldsymbol{q}}_3 h_2 \left[\frac{\partial a_3}{\partial q_2} - h_3 a_2 \frac{\partial}{\partial q_3} \left(\frac{1}{h_2} \right) \right] \\ & + \hat{\boldsymbol{q}}_3 \hat{\boldsymbol{q}}_1 h_3 \left[\frac{\partial a_1}{\partial q_3} - h_1 a_3 \frac{\partial}{\partial q_1} \left(\frac{1}{h_3} \right) \right] \\ & + \hat{\boldsymbol{q}}_3 \hat{\boldsymbol{q}}_2 h_3 \left[\frac{\partial a_2}{\partial q_3} - h_2 a_3 \frac{\partial}{\partial q_2} \left(\frac{1}{h_3} \right) \right] \\ & + \hat{\boldsymbol{q}}_3 \hat{\boldsymbol{q}}_3 h_3 \left[\frac{\partial a_3}{\partial q_3} + h_1 a_1 \frac{\partial}{\partial q_1} \left(\frac{1}{h_3} \right) + h_2 a_2 \frac{\partial}{\partial q_2} \left(\frac{1}{h_3} \right) \right]. \end{aligned} \tag{1.252}$$

Last, the divergence of a second-order tensor $\boldsymbol{A}$ is then

$$\begin{aligned}
\nabla\cdot\boldsymbol{A} = \hat{\boldsymbol{q}}_1 \Bigg\{ & h_1h_2h_3\left[\frac{\partial}{\partial q_1}\left(\frac{A_{11}}{h_2h_3}\right)+\frac{\partial}{\partial q_2}\left(\frac{A_{21}}{h_1h_3}\right)+\frac{\partial}{\partial q_3}\left(\frac{A_{31}}{h_1h_2}\right)\right] \\
& + \cancel{h_1h_1A_{11}\frac{\partial}{\partial q_1}\left(\frac{1}{h_1}\right)} + h_1h_2A_{12}\frac{\partial}{\partial q_2}\left(\frac{1}{h_1}\right) + h_1h_3A_{13}\frac{\partial}{\partial q_3}\left(\frac{1}{h_1}\right) \\
& - \cancel{h_1h_1A_{11}\frac{\partial}{\partial q_1}\left(\frac{1}{h_1}\right)} - h_1h_2A_{22}\frac{\partial}{\partial q_1}\left(\frac{1}{h_2}\right) - h_1h_3A_{33}\frac{\partial}{\partial q_1}\left(\frac{1}{h_3}\right)\Bigg\} \\
& + \hat{\boldsymbol{q}}_2\left\{h_1h_2h_3\left[\frac{\partial}{\partial q_1}\left(\frac{A_{12}}{h_2h_3}\right)+\frac{\partial}{\partial q_2}\left(\frac{A_{22}}{h_1h_3}\right)+\frac{\partial}{\partial q_3}\left(\frac{A_{32}}{h_1h_2}\right)\right]\right. \\
& + h_2h_1A_{21}\frac{\partial}{\partial q_1}\left(\frac{1}{h_2}\right) + \cancel{h_2h_2A_{22}\frac{\partial}{\partial q_2}\left(\frac{1}{h_2}\right)} + h_2h_3A_{23}\frac{\partial}{\partial q_3}\left(\frac{1}{h_2}\right) \\
& \left. - h_2h_1A_{11}\frac{\partial}{\partial q_2}\left(\frac{1}{h_1}\right) - \cancel{h_2h_2A_{22}\frac{\partial}{\partial q_2}\left(\frac{1}{h_2}\right)} - h_2h_3A_{33}\frac{\partial}{\partial q_2}\left(\frac{1}{h_3}\right)\right\} \\
& + \hat{\boldsymbol{q}}_3\left\{h_1h_2h_3\left[\frac{\partial}{\partial q_1}\left(\frac{A_{13}}{h_2h_3}\right)+\frac{\partial}{\partial q_2}\left(\frac{A_{23}}{h_1h_3}\right)+\frac{\partial}{\partial q_3}\left(\frac{A_{33}}{h_1h_2}\right)\right]\right. \\
& + h_3h_1A_{31}\frac{\partial}{\partial q_1}\left(\frac{1}{h_3}\right) + h_3h_2A_{32}\frac{\partial}{\partial q_2}\left(\frac{1}{h_3}\right) + \cancel{h_3h_3A_{33}\frac{\partial}{\partial q_3}\left(\frac{1}{h_3}\right)} \\
& \left. - h_3h_1A_{11}\frac{\partial}{\partial q_3}\left(\frac{1}{h_1}\right) - h_3h_2A_{22}\frac{\partial}{\partial q_3}\left(\frac{1}{h_2}\right) - \cancel{h_3h_3A_{33}\frac{\partial}{\partial q_3}\left(\frac{1}{h_3}\right)}\right\}.
\end{aligned} \tag{1.253}$$

This can be compared to the same expression given in Cartesian coordinates

$$\begin{aligned}
\nabla\cdot\boldsymbol{A} = {} & \hat{\boldsymbol{x}}_1\left(\frac{\partial A_{11}}{\partial x_1}+\frac{\partial A_{21}}{\partial x_2}+\frac{\partial A_{31}}{\partial x_3}\right)+\hat{\boldsymbol{x}}_2\left(\frac{\partial A_{12}}{\partial x_1}+\frac{\partial A_{22}}{\partial x_2}+\frac{\partial A_{32}}{\partial x_3}\right) \\
& + \hat{\boldsymbol{x}}_3\left(\frac{\partial A_{13}}{\partial x_1}+\frac{\partial A_{23}}{\partial x_2}+\frac{\partial A_{33}}{\partial x_3}\right).
\end{aligned} \tag{1.254}$$

Using the above, the most pertinent expressions for the special cases of cylindrical and spherical coordinates follow.

Cylindrical Coordinates

In cylindrical coordinates $(q_1, q_2, q_3) \,\widehat{=}\, (r, \theta, z)$ with unit vectors $\hat{\boldsymbol{r}}$ (radial direction), $\hat{\boldsymbol{\theta}}$ (circumferential direction around the z axis), and $\hat{z}$ (axial direction) that are orthogonal to each other at each point in space, the metrical coefficients are

$$\frac{1}{h_1} = 1, \qquad \frac{1}{h_2} = r \quad \text{and} \quad \frac{1}{h_3} = 1. \tag{1.255}$$

This simply says that distance in the θ direction goes as θr. The mapping of the Cartesian base vectors into the cylindrical-coordinate base vectors is performed using the matrix multiplication

$$\begin{bmatrix} \hat{\boldsymbol{r}} \\ \hat{\boldsymbol{\theta}} \\ \hat{\boldsymbol{z}} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \\ \hat{\boldsymbol{z}} \end{bmatrix}. \tag{1.256}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

Some further trigonometry gives $r=\sqrt{x^2+y^2}$, $\theta=\tan^{-1}(y/x)$, and $z=z$ as well as $x=r\cos\theta$ and $y=r\sin\theta$.

The nonzero derivatives of the base vectors are given by Eqs (1.238) and (1.239) to be

$$\frac{\partial\hat{\boldsymbol{r}}}{\partial\theta}=\hat{\boldsymbol{\theta}} \quad \text{and} \quad \frac{\partial\hat{\boldsymbol{\theta}}}{\partial\theta}=-\hat{\boldsymbol{r}}. \tag{1.257}$$

Consider three fields: $\psi(\boldsymbol{r})$ a scalar, $\boldsymbol{a}(\boldsymbol{r})=(a_r,a_\theta,a_z)$ a vector, and

$$\boldsymbol{A}(\boldsymbol{r})=\begin{pmatrix} A_{rr} & A_{r\theta} & A_{rz} \\ A_{\theta r} & A_{\theta\theta} & A_{\theta z} \\ A_{zr} & A_{z\theta} & A_{zz} \end{pmatrix} \tag{1.258}$$

a second-order tensor. The various standard operations in cylindrical coordinates involving ∇ acting on the scalar and vector fields are

$$\nabla\psi=\hat{\boldsymbol{r}}\frac{\partial\psi}{\partial r}+\hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial\psi}{\partial\theta}+\hat{\boldsymbol{z}}\frac{\partial\psi}{\partial z} \tag{1.259}$$

$$\nabla^2\psi=\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\partial^2\psi}{\partial z^2} \tag{1.260}$$

and

$$\nabla\cdot\boldsymbol{a}=\frac{1}{r}\frac{\partial}{\partial r}(ra_r)+\frac{1}{r}\frac{\partial a_\theta}{\partial\theta}+\frac{\partial a_z}{\partial z} \tag{1.261}$$

$$\nabla\times\boldsymbol{a}=\hat{\boldsymbol{r}}\left(\frac{1}{r}\frac{\partial a_z}{\partial\theta}-\frac{\partial a_\theta}{\partial z}\right)+\hat{\boldsymbol{\theta}}\left(\frac{\partial a_r}{\partial z}-\frac{\partial a_z}{\partial r}\right)+\hat{\boldsymbol{z}}\left(\frac{1}{r}\frac{\partial}{\partial r}(ra_\theta)-\frac{1}{r}\frac{\partial a_r}{\partial\theta}\right) \tag{1.262}$$

$$\nabla^2\boldsymbol{a}=\hat{\boldsymbol{r}}\left(\nabla^2a_r-\frac{2}{r^2}\frac{\partial a_\theta}{\partial\theta}-\frac{a_r}{r^2}\right)+\hat{\boldsymbol{\theta}}\left(\nabla^2a_\theta+\frac{2}{r^2}\frac{\partial a_r}{\partial\theta}-\frac{a_\theta}{r^2}\right)+\hat{\boldsymbol{z}}\nabla^2u_z. \tag{1.263}$$

In this last expression, the Laplacian operator ∇^2 acting on the three scalar components of the vector $\boldsymbol{a}$ is given by Eq. (1.260). The two most common tensorial operations we will encounter are

$$\begin{aligned}\nabla\boldsymbol{a}=&\,\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}\frac{\partial a_r}{\partial r}+\hat{\boldsymbol{r}}\hat{\boldsymbol{\theta}}\frac{\partial a_\theta}{\partial r}+\hat{\boldsymbol{r}}\hat{\boldsymbol{z}}\frac{\partial a_z}{\partial r}\\ &+\hat{\boldsymbol{\theta}}\hat{\boldsymbol{r}}\frac{1}{r}\left(\frac{\partial a_r}{\partial\theta}-a_\theta\right)+\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}\frac{1}{r}\left(\frac{\partial a_\theta}{\partial\theta}+a_r\right)+\hat{\boldsymbol{\theta}}\hat{\boldsymbol{z}}\frac{1}{r}\frac{\partial a_z}{\partial\theta}\\ &+\hat{\boldsymbol{z}}\hat{\boldsymbol{r}}\frac{\partial a_r}{\partial z}+\hat{\boldsymbol{z}}\hat{\boldsymbol{\theta}}\frac{\partial a_\theta}{\partial z}+\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\frac{\partial a_z}{\partial z}\end{aligned} \tag{1.264}$$

and

$$\begin{aligned}\nabla\cdot\boldsymbol{A}=&\,\hat{\boldsymbol{r}}\left[\frac{1}{r}\frac{\partial}{\partial r}(rA_{rr})+\frac{1}{r}\frac{\partial A_{\theta r}}{\partial\theta}+\frac{\partial A_{zr}}{\partial z}-\frac{A_{\theta\theta}}{r}\right]\\ &+\hat{\boldsymbol{\theta}}\left[\frac{1}{r}\frac{\partial}{\partial r}(rA_{r\theta})+\frac{1}{r}\frac{\partial A_{\theta\theta}}{\partial\theta}+\frac{\partial A_{z\theta}}{\partial z}+\frac{A_{\theta r}}{r}\right]\\ &+\hat{\boldsymbol{z}}\left[\frac{1}{r}\frac{\partial}{\partial r}(rA_{rz})+\frac{1}{r}\frac{\partial A_{\theta z}}{\partial\theta}+\frac{\partial A_{zz}}{\partial z}\right].\end{aligned} \tag{1.265}$$

Spherical Coordinates

In spherical coordinates (r, θ, ϕ), with θ now measuring latitude down from a "z axis" of revolution and ϕ measuring longitude around the z axis and with unit vectors $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ that are orthogonal to each other at each point in space, the metrical coefficients are

$$\frac{1}{h_1} = 1, \quad \frac{1}{h_2} = r, \quad \text{and} \quad \frac{1}{h_3} = r\sin\theta. \tag{1.266}$$

This says that at each latitude θ coming down from the z axis, distance in the longitudinal direction around the z axis goes as $\phi r \sin\theta$. The mapping of the Cartesian base vectors into the spherical-coordinate base vectors is performed using the matrix multiplication

$$\begin{bmatrix} \hat{\boldsymbol{r}} \\ \hat{\boldsymbol{\theta}} \\ \hat{\boldsymbol{\phi}} \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \\ \hat{\boldsymbol{z}} \end{bmatrix}. \tag{1.267}$$

Some further trigonometry gives $r = \sqrt{x^2 + y^2 + z^2}$, $\theta = \cos^{-1}\left(z/\sqrt{x^2 + y^2 + z^2}\right)$, and $\phi = \tan^{-1}(y/x)$ as well as $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\theta$.

Of the nine possible derivatives of the base vectors, Eqs (1.238) and (1.239) give that five are nonzero

$$\frac{\partial \hat{\boldsymbol{r}}}{\partial \theta} = \hat{\boldsymbol{\theta}} \quad \text{and} \quad \frac{\partial \hat{\boldsymbol{\theta}}}{\partial \theta} = -\hat{\boldsymbol{r}} \tag{1.268}$$

as well as

$$\frac{\partial \hat{\boldsymbol{r}}}{\partial \phi} = \hat{\boldsymbol{\phi}}\sin\theta, \quad \frac{\partial \hat{\boldsymbol{\theta}}}{\partial \phi} = \hat{\boldsymbol{\phi}}\cos\theta, \quad \text{and} \quad \frac{\partial \hat{\boldsymbol{\phi}}}{\partial \phi} = -\hat{\boldsymbol{r}}\sin\theta - \hat{\boldsymbol{\theta}}\cos\theta. \tag{1.269}$$

Again consider three fields: $\psi(\boldsymbol{r})$ a scalar, $\boldsymbol{a}(\boldsymbol{r}) = (a_r, a_\theta, a_\phi)$ a vector, and

$$\boldsymbol{A}(\boldsymbol{r}) = \begin{pmatrix} A_{rr} & A_{r\theta} & A_{r\phi} \\ A_{\theta r} & A_{\theta\theta} & A_{\theta\phi} \\ A_{\phi r} & A_{\phi\theta} & A_{\phi\phi} \end{pmatrix} \tag{1.270}$$

a second-order tensor. The standard operations in spherical coordinates involving ∇ and the scalar and vector fields are

$$\nabla\psi = \hat{\boldsymbol{r}}\frac{\partial\psi}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial\psi}{\partial\theta} + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi} \tag{1.271}$$

$$\nabla^2\psi = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right] \tag{1.272}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

and

$$\nabla \cdot \boldsymbol{a} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 a_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\, a_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial a_\phi}{\partial\phi} \tag{1.273}$$

$$\begin{aligned}\nabla \times \boldsymbol{a} = {} & \hat{\boldsymbol{r}}\frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta\, a_\phi\right) - \frac{\partial a_\theta}{\partial\phi}\right) + \hat{\boldsymbol{\theta}}\frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial a_r}{\partial\phi} - \frac{\partial}{\partial r}\left(r a_\phi\right)\right) \\ & + \hat{\boldsymbol{\phi}}\frac{1}{r}\left(\frac{\partial}{\partial r}\left(r a_\theta\right) - \frac{\partial a_r}{\partial\theta}\right)\end{aligned} \tag{1.274}$$

$$\begin{aligned}\nabla^2 \boldsymbol{a} = {} & \hat{\boldsymbol{r}}\left[\nabla^2 a_r - \frac{2}{r^2}\left(a_r + \frac{\partial a_\theta}{\partial\theta} + \frac{\cos\theta}{\sin\theta}a_\theta - \frac{1}{\sin\theta}\frac{\partial a_\phi}{\partial\phi}\right)\right] \\ & + \hat{\boldsymbol{\theta}}\left[\nabla^2 a_\theta + \frac{1}{r^2}\left(2\frac{\partial a_r}{\partial\theta} - \frac{a_\theta}{\sin^2\theta} - \frac{2\cos\theta}{\sin^2\theta}\frac{\partial a_\phi}{\partial\phi}\right)\right] \\ & + \hat{\boldsymbol{\phi}}\left[\nabla^2 a_\phi + \frac{1}{r^2}\left(\frac{2}{\sin\theta}\frac{\partial a_r}{\partial\phi} + \frac{2\cos\theta}{\sin^2\theta}\frac{\partial a_\theta}{\partial\phi} - \frac{a_\phi}{\sin^2\theta}\right)\right].\end{aligned} \tag{1.275}$$

In this last expression, the Laplacian operator ∇^2 acting on the three scalar components of the vector $\boldsymbol{a}$ is given by Eq. (1.272). The two most common tensorial operations involving ∇ are

$$\begin{aligned}\nabla \boldsymbol{a} = {} & \hat{\boldsymbol{r}}\hat{\boldsymbol{r}}\frac{\partial a_r}{\partial r} + \hat{\boldsymbol{r}}\hat{\boldsymbol{\theta}}\frac{\partial a_\theta}{\partial r} + \hat{\boldsymbol{r}}\hat{\boldsymbol{\phi}}\frac{\partial a_\phi}{\partial r} \\ & + \hat{\boldsymbol{\theta}}\hat{\boldsymbol{r}}\frac{1}{r}\left(\frac{\partial a_r}{\partial\theta} - a_\theta\right) + \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}\frac{1}{r}\left(\frac{\partial a_\theta}{\partial\theta} + a_r\right) + \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\phi}}\frac{1}{r}\frac{\partial a_\phi}{\partial\theta} \\ & + \hat{\boldsymbol{\phi}}\hat{\boldsymbol{r}}\frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial a_r}{\partial\phi} - a_\phi\right) + \hat{\boldsymbol{\phi}}\hat{\boldsymbol{\theta}}\frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial a_\theta}{\partial\phi} - \frac{\cos\theta}{\sin\theta}a_\phi\right) \\ & + \hat{\boldsymbol{\phi}}\hat{\boldsymbol{\phi}}\frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial a_\phi}{\partial\phi} + a_r + \frac{\cos\theta}{\sin\theta}a_\theta\right)\end{aligned} \tag{1.276}$$

and

$$\begin{aligned}\nabla \cdot \boldsymbol{A} = {} & \hat{\boldsymbol{r}}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 A_{rr}\right) + \frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta A_{\theta r}\right) + \frac{\partial A_{\phi r}}{\partial\phi}\right) - \frac{\left(A_{\theta\theta} + A_{\phi\phi}\right)}{r}\right] \\ & + \hat{\boldsymbol{\theta}}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 A_{r\theta}\right) + \frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta A_{\theta\theta}\right) + \frac{\partial A_{\phi\theta}}{\partial\phi}\right) + \frac{A_{\theta r}}{r} - \frac{\cos\theta}{r\sin\theta}A_{\phi\phi}\right] \\ & + \hat{\boldsymbol{\phi}}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 A_{r\phi}\right) + \frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta A_{\theta\phi}\right) + \frac{\partial A_{\phi\phi}}{\partial\phi}\right) + \frac{A_{\phi r}}{r} - \frac{\cos\theta}{r\sin\theta}A_{\theta\phi}\right].\end{aligned} \tag{1.277}$$

1.9 The Dirac Delta Function

Throughout our development and implementation of continuum physics, we need to represent fields that are highly concentrated at a single point in space (or time, if considering time functions). The *Dirac delta* function $\delta(x - x_1)$ is used to represent a field highly concentrated at the point $x = x_1$ and is loosely, though insufficiently, defined as

$$\delta(x-x_1) = \begin{cases} \infty & \text{if } x = x_1 \\ 0 & \text{if } x \neq x_1 \end{cases} \tag{1.278}$$

and in such a way that if $a < x_1 < b$, then

$$\int_a^b \delta(x-x_1)\,\mathrm{d}x = 1 \quad \text{(no units)}. \tag{1.279}$$

In some sense, we have that when $x = x_1$, then $\delta(0) = \mathrm{d}x^{-1}$ so that the *integral property of the Dirac delta function* of Eq. (1.279) is satisfied (indeed, when working numerically on discretized domains, this is one way to define the Dirac delta). Due to this integral property, we also have

$$\boxed{\int_a^b f(x)\,\delta(x-x_1)\,\mathrm{d}x = f(x_1),} \tag{1.280}$$

which is called the *sifting property*. In the integration process, the Dirac delta function samples f, where the argument of the Dirac goes to zero. Note as well that from Eq. (1.279) (or equivalently the sifting property), we necessarily have that $\delta(x)$ has physical of units of x^{-1} whatever the physical units of x are. This is important to remember in physics applications.

Because Eq. (1.278) is not a sufficient definition for a well-behaved differentiable function, we better define the Dirac delta function $\delta(x-x_1)$ as the limit of well-defined functions such as

$$\delta(x-x_1) = \lim_{\sigma\to 0} \frac{S(x-x_1+\sigma) - S(x-x_1-\sigma)}{2\sigma} \tag{1.281}$$

$$= \lim_{\sigma\to 0} \frac{\sin\left[(x-x_1)/\sigma\right]}{\pi(x-x_1)} \tag{1.282}$$

$$= \lim_{\sigma\to 0} \frac{1}{\sigma\pi\left[1+(x-x_1)^2/\sigma^2\right]} \tag{1.283}$$

$$= \lim_{\sigma\to 0} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-x_1)^2}{2\sigma^2}\right), \tag{1.284}$$

where in the first example here, $S(x)$ is the unit step function defined to be 0 for $x < 0$ and 1 for $x > 0$. Each of these expressions for $\delta(x-x_1)$ satisfies the integral constraint of Eq. (1.279) for any value of σ. In the limit as the parameter σ (which has the same units as x and characterizes the width of the function) goes to zero, these also satisfy the sifting property of Eq. (1.280). The second example here is the scaled "sinc" function and will be shown to satisfy the required integral properties in Chapter 11 on contour integration. The last example is the familiar Gaussian function with standard deviation σ and will be shown to satisfy the required integral properties in Chapter 10 on Fourier analysis. We will use the Dirac to represent highly concentrated fields in nature; however, in each application

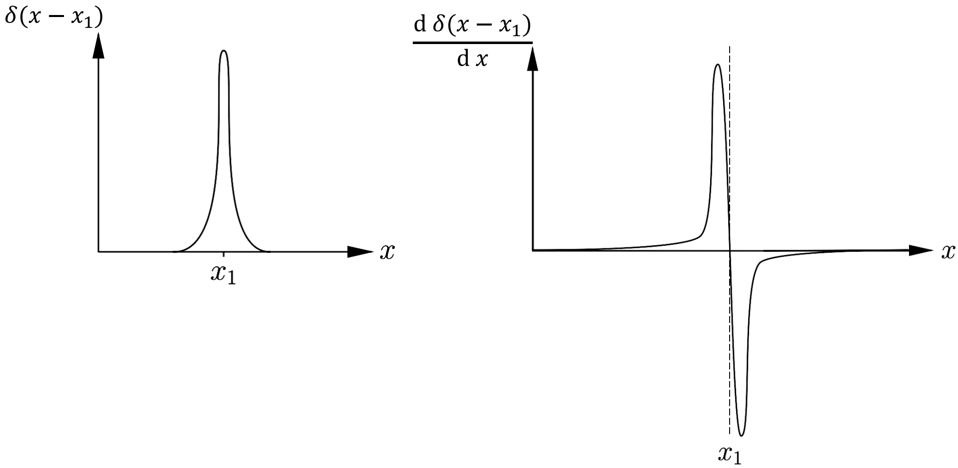

Figure 1.15 The Dirac delta and its derivative for small σ in Eqs (1.284) and (1.285).

to a physics problem, the Dirac delta function will be integrated over. Due to the defining integrals of Eqs (1.279) and (1.280), such integration leads to well-behaved finite results despite the fact that the Dirac function becomes very large when its argument is zero and as $\sigma \to 0$ (small but finite) in Eqs (1.281)–(1.284).

By representing the Dirac as a limit of a well-behaved function, one can take derivatives of $\delta(x - x_1)$. For example, for the Gaussian representation of $\delta(x - x_1)$, we have

$$\frac{\mathrm{d}\,\delta(x - x_1)}{\mathrm{d}x} = \lim_{\sigma \to 0} -\frac{(x - x_1)}{\sigma^3 \sqrt{2\pi}} \exp\left(\frac{-(x - x_1)^2}{2\sigma^2}\right). \tag{1.285}$$

Visually, one can see the effect of taking the derivative of a Dirac delta by using Eqs (1.284) and (1.285) as shown in Fig. 1.15. One can also understand the derivative of the Dirac delta through the usual definition of the derivative, which is equivalent to taking the derivative of Eq. (1.281)

$$\frac{\mathrm{d}\delta(x - x_1)}{\mathrm{d}x} = \lim_{\sigma \to 0} \frac{\delta(x - x_1 + \sigma) - \delta(x - x_1 - \sigma)}{2\sigma}, \tag{1.286}$$

that is, as the sum of two Dirac functions of opposite sign that approach each other, also as depicted in Fig. 1.15.

To use the derivative of the Dirac delta function, note that if $a < x_1 < b$,

$$\int_a^b f(x) \frac{\mathrm{d}\,\delta(x - x_1)}{\mathrm{d}x}\,\mathrm{d}x = \int_a^b \left\{ \frac{\mathrm{d}}{\mathrm{d}x}\Big[f(x)\,\delta(x - x_1)\Big] - \frac{\mathrm{d}f(x)}{\mathrm{d}x}\delta(x - x_1) \right\} \mathrm{d}x \tag{1.287}$$

$$= -\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x = x_1} \tag{1.288}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

because $\delta(a-x_1)=0$ and $\delta(b-x_1)=0$. Equation (1.288) is called the *derivative-sifting property* and generalizes to an arbitrary number of derivatives n as

$$\int_a^b f(x)\,\frac{\mathrm{d}^n\delta(x-x_1)}{\mathrm{d}x^n}\,\mathrm{d}x = (-1)^n \left.\frac{\mathrm{d}^n f(x)}{\mathrm{d}x^n}\right|_{x=x_1}. \tag{1.289}$$

So taking the derivative of a Dirac delta function is legitimate at least when one then integrates with it, which you will see is always done in all applications. In Chapter 2, the Dirac delta is used to represent the position of single atoms. In this same context, the derivative of the Dirac is used to represent electric dipoles, in which a concentration of discrete positive charge is located a small distance σ away from a concentration of discrete negative charge.

We can define the *unit step function* $S(x-x_1)$ as the integral of the Dirac delta function

$$S(x-x_1) = \int_{-\infty}^{x} \delta(x_0-x_1)\,\mathrm{d}x_0 \,\widehat{=}\, \begin{cases} 1 & x \geq x_1 \\ 0 & x < x_1 \end{cases}. \tag{1.290}$$

Note that the 1 here is unitless regardless of the physical units of x. The *ramp function* $R(x-x_1)$ is similarly defined as the integral of the step function

$$R(x-x_1) = (x-x_1)S(x-x_1) = \int_{-\infty}^{x} \mathrm{d}x_0 \int_{-\infty}^{x_0} \delta(x_2-x_1)\,\mathrm{d}x_2$$
$$= \int_{-\infty}^{x} \mathrm{d}x_0\, S(x_0-x_1) \,\widehat{=}\, \begin{cases} x-x_1 & x \geq x_1 \\ 0 & x < x_1 \end{cases}. \tag{1.291}$$

Thus we also have that

$$\boxed{\delta(x-x_1) = \frac{\mathrm{d}}{\mathrm{d}x}S(x-x_1) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}R(x-x_1).} \tag{1.292}$$

We will use these results for the step and ramp functions in the development of our elastodynamic Green's tensor in Chapter 12.

If we make the substitution of variables that $x \to ax'$ where a is some scalar, then $\mathrm{d}x = a\mathrm{d}x'$ and

$$\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1 = \int_{-\infty}^{\infty} \delta(ax')\,a\mathrm{d}x'. \tag{1.293}$$

If a is negative, the integral would be from $+\infty$ to $-\infty$ which, upon using that $\int_{+\infty}^{-\infty}\mathrm{d}x' = -\int_{-\infty}^{+\infty}\mathrm{d}x'$, yields the same result of 1. So we can conclude that

$$\delta(ax) = \frac{\delta(x)}{|a|}, \tag{1.294}$$

which is important to remember in some applications. For example, if a Dirac function is moving about as a wave response with wave speed c, this can be expressed $\delta(t-r/c) = \delta(r/c-t) = \delta((r-ct)/c) = c\delta(r-ct)$, which is good to know when comparing solutions of the wave equation obtained using different approaches.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

One can further generalize to consider a Dirac delta function that is a function of another function, say $g(x)$, that has zeroes and ask about the nature of the integral

$$I = \int_{-\infty}^{\infty} \mathrm{d}x\, \delta[g(x)] f(x). \tag{1.295}$$

To treat this integral, we note that where $g(x) \neq 0$ the Dirac function is zero. So there is only contribution to the integral at the places where $x \to x_i$ where the x_i are the zeroes of $g(x)$ (i.e., the places where $g(x_i) = 0$). As $x \to x_i$ we can represent $g(x)$ as the lead term of the Taylor expansion of $g(x)$, which is $(x - x_i)g'(x_i)$ where $g'(x_i) \,\widehat{=}\, \mathrm{d}g(x)/\mathrm{d}x|_{x=x_i}$. Thus, we can write

$$\delta[g(x)] = \sum_i \delta[(x - x_i)g'(x_i)] = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|} \tag{1.296}$$

so that we have

$$I = \int_{-\infty}^{\infty} \mathrm{d}x\, \delta[g(x)] f(x) = \sum_i \frac{f(x_i)}{|g'(x_i)|} \tag{1.297}$$

where the sum is over all the zeroes x_i of $g(x)$ found within the domain of integration.

To represent a field concentrated at a point $\boldsymbol{r}_1$ in 3D space using Cartesian coordinates, we use the notation

$$\delta(\boldsymbol{r} - \boldsymbol{r}_1) \,\widehat{=}\, \delta(x - x_1)\, \delta(y - y_1)\, \delta(z - z_1). \tag{1.298}$$

Note that $\delta(\boldsymbol{r} - \boldsymbol{r}_1)$ has units of inverse length cubed, because $\delta(\psi)$ with ψ some scalar, has units of ψ^{-1} as Eq. (1.279) or Eq. (1.294) makes clear. In cylindrical coordinates (r, θ, z), we employ the notation

$$\delta(\boldsymbol{r} - \boldsymbol{r}_1) = \delta(r - r_1)\delta[r(\theta - \theta_1)]\delta(z - z_1) = \delta(r - r_1)\frac{\delta(\theta - \theta_1)}{r}\delta(z - z_1), \tag{1.299}$$

where we used the scaling property of Eq. (1.294). In this cylindrical-coordinate notation, the integrated result over the 3D whole space Ω_∞ is

$$\int_{\Omega_\infty} \delta(\boldsymbol{r} - \boldsymbol{r}_1)\, \mathrm{d}V = \int_0^{\infty} \mathrm{d}r \int_0^{2\pi} r\mathrm{d}\theta \int_{-\infty}^{\infty} \mathrm{d}z\, \delta(r - r_1)\frac{\delta(\theta - \theta_1)}{r}\delta(z - z_1) = 1, \tag{1.300}$$

where we used that $\mathrm{d}V = (\mathrm{d}r)(r\mathrm{d}\theta)(\mathrm{d}z)$ in cylindrical coordinates. Similarly, in spherical coordinates (r, θ, ϕ), we use

$$\delta(\boldsymbol{r} - \boldsymbol{r}_1) = \delta(r - r_1)\delta[r(\theta - \theta_1)]\delta[r \sin\theta(\phi - \phi_1)] = \frac{\delta(r - r_1)\delta(\theta - \theta_1)\delta(\phi - \phi_1)}{r^2 \sin\theta} \tag{1.301}$$

to represent a field concentrated at a point. Employing $\mathrm{d}V = (\mathrm{d}r)(r\mathrm{d}\theta)(r\sin\theta\mathrm{d}\phi)$ then gives again

$$\int_{\Omega_\infty} \delta(\boldsymbol{r}-\boldsymbol{r}_1)\,\mathrm{d}V = \int_0^\infty \mathrm{d}r \int_0^\pi r\mathrm{d}\theta \int_0^{2\pi} r\sin\theta\mathrm{d}\phi\,\frac{\delta(r-r_1)\delta(\theta-\theta_1)\delta(\phi-\phi_1)}{r^2\sin\theta} = 1 \tag{1.302}$$

as required.

In 3D space, we then have the sifting property for a point $\boldsymbol{r}_1$ lying within a volumetric region Ω

$$\int_\Omega \delta(\boldsymbol{r}-\boldsymbol{r}_1)\psi(\boldsymbol{r})\,\mathrm{d}V = \psi(\boldsymbol{r}_1), \tag{1.303}$$

where the field ψ can be a scalar, vector, or tensor of any order. We also have the 3D version of the derivative sifting property

$$\int_\Omega [\nabla\delta(\boldsymbol{r}-\boldsymbol{r}_1)]\,\psi(\boldsymbol{r})\,\mathrm{d}V = -\,\nabla\psi|_{\boldsymbol{r}=\boldsymbol{r}_1} \tag{1.304}$$

and for multiple applications of the gradient operator

$$\int_\Omega [{}_n\nabla\delta(\boldsymbol{r}-\boldsymbol{r}_1)]\,\psi(\boldsymbol{r})\,\mathrm{d}V = (-1)^n\,{}_n\nabla\psi|_{\boldsymbol{r}=\boldsymbol{r}_1}, \tag{1.305}$$

where ${}_n\nabla = \nabla\nabla\ldots\nabla$ represents n successive ∇ operations.

1.10 Exercises

1. Through direct calculation of the derivatives, demonstrate that

$$\nabla\cdot(\nabla\times\boldsymbol{a}) = 0 \tag{1.306}$$

$$\nabla\times(\nabla\psi) = 0. \tag{1.307}$$

Now, prove these same identities, again through direct calculation, using the Levi–Civita alternating third-order tensor.

2. Using the method demonstrated in Section 1.7, in which the tensorial expressions are first expressed in Cartesian coordinates, derivatives between the scalar components carried out using the usual product rule and dot products performed between base vectors prior to returning to the bold-face representation valid for all coordinate systems, prove all eleven of the identities in the list of Eqs (1.98)–(1.115) that do not involve a curl operation. For an even greater challenge, also prove the identities involving the curl by using the Levi–Civita tensor.

3. With $\boldsymbol{I} = \delta_{ij}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j$ being the identity tensor and $\boldsymbol{r} = x_i\hat{\boldsymbol{x}}_i$ the position vector in Cartesian coordinates, prove

$$\nabla\boldsymbol{r} = \boldsymbol{I} \tag{1.308}$$

$$\nabla\cdot\boldsymbol{r} = 3. \tag{1.309}$$

4. If $\boldsymbol{a}$ is some spatially variable vector field in a region Ω and $\boldsymbol{n}$ is the outward normal to the closed surface $\partial\Omega$ that surrounds Ω, show that

$$\int_{\partial\Omega} \boldsymbol{a}\boldsymbol{n}\, \mathrm{d}S = \int_{\Omega} (\nabla \boldsymbol{a})^T\, \mathrm{d}V. \tag{1.310}$$

5. *Prove the fundamental theorem of 3D calculus*: If $\boldsymbol{a}$ is any spatially variable and differentiable vector field $\boldsymbol{a}(x, y, z) = a_x(x, y, z)\hat{\boldsymbol{x}} + a_y(x, y, z)\hat{\boldsymbol{y}} + a_z(x, y, z)\hat{\boldsymbol{z}}$, use the fundamental theorem of 1D calculus to show that if region Ω is a cube with sides of length L then

$$\int_{\Omega} \nabla \boldsymbol{a}\, \mathrm{d}^3\boldsymbol{r} \widehat{=} \int_0^L \mathrm{d}x \int_0^L \mathrm{d}y \int_0^L \mathrm{d}z\, \nabla \boldsymbol{a}(x, y, z) \tag{1.311}$$

$$\begin{aligned} &= \int_0^L \mathrm{d}y \int_0^L \mathrm{d}z\, \hat{\boldsymbol{x}}\boldsymbol{a}(L, y, z) - \int_0^L \mathrm{d}y \int_0^L \mathrm{d}z\, \hat{\boldsymbol{x}}\boldsymbol{a}(0, y, z) \\ &\quad + \int_0^L \mathrm{d}x \int_0^L \mathrm{d}z\, \hat{\boldsymbol{y}}\boldsymbol{a}(x, L, z) - \int_0^L \mathrm{d}x \int_0^L \mathrm{d}z\, \hat{\boldsymbol{y}}\boldsymbol{a}(x, 0, z) \\ &\quad + \int_0^L \mathrm{d}x \int_0^L \mathrm{d}y\, \hat{\boldsymbol{z}}\boldsymbol{a}(x, y, L) - \int_0^L \mathrm{d}x \int_0^L \mathrm{d}y\, \hat{\boldsymbol{z}}\boldsymbol{a}(x, y, 0). \end{aligned} \tag{1.312}$$

Then show that the six surface integrals on the right-hand side here are the contributions from each of the six cube faces coming from the surface integral

$$\int_{\partial\Omega} \boldsymbol{n}\boldsymbol{a}\, \mathrm{d}^2\boldsymbol{r}. \tag{1.313}$$

You thus obtain the fundamental theorem of 3D calculus $\int_{\Omega} \nabla \boldsymbol{a}\, \mathrm{d}^3\boldsymbol{r} = \int_{\partial\Omega} \boldsymbol{n}\boldsymbol{a}\, \mathrm{d}^2\boldsymbol{r}$ using the fundamental theorem of 1D calculus for the case of a cubic integration domain. For an arbitrarily shaped region Ω, you fill the region with tiny (approaching infinitesimal) cubes in a cubic packing. The sum of the volume integrals of $\nabla \boldsymbol{a}$ for each tiny cube adds up to the volume integral of $\nabla \boldsymbol{a}$ over all of Ω. For adjacent tiny cubes that share the same surface, the normal for each cube is oppositely directed and have surface integrals of $\boldsymbol{n}\boldsymbol{a}$ that cancel when summing over the cubes except for the surfaces that are coincident with $\partial\Omega$. The fundamental theorem of 3D calculus is thus proven for arbitrarily shaped regions.

6. Starting from each of the nine components represented by the second-order tensor integral identity

$$\int_{\Omega} \nabla \boldsymbol{a}\, \mathrm{d}^3\boldsymbol{r} = \int_{\partial\Omega} \boldsymbol{n}\boldsymbol{a}\, \mathrm{d}^2\boldsymbol{r}, \tag{1.314}$$

where $\boldsymbol{a}$ is some spatially variable vector field, work in Cartesian coordinates to demonstrate the divergence theorem

$$\int_{\Omega} \nabla \cdot \boldsymbol{a}\, \mathrm{d}^3\boldsymbol{r} = \int_{\partial\Omega} \boldsymbol{n} \cdot \boldsymbol{a}\, \mathrm{d}^2\boldsymbol{r}. \tag{1.315}$$

7. *Prove Green's theorem*: Green's theorem is the statement that on the (x, y) plane, two differentiable functions $P(x, y)$ and $Q(x, y)$ satisfy

$$\int_S \left(\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y} \right) \mathrm{d}x\, \mathrm{d}y = \oint_\Gamma \left[P(x,y)\, \mathrm{d}x + Q(x,y)\, \mathrm{d}y \right], \tag{1.316}$$

where S is any surface on the (x, y) plane that is bounded by the closed contour Γ and with the sense of the contour integral being counterclockwise. This can be proven rather trivially for the case of the rectangular surface shown in Fig. 1.16. To do so, simply integrate the left-hand side of Eq. (1.316) over the rectangle and use the 1D fundamental theorem of calculus to obtain

$$\int_{y_1}^{y_2} \mathrm{d}y \int_{x_1}^{x_2} \mathrm{d}x \frac{\partial Q(x,y)}{\partial x} = \int_{y_1}^{y_2} \mathrm{d}y \left[Q(x_2, y) - Q(x_1, y) \right] \tag{1.317}$$

$$-\int_{x_1}^{x_2} \mathrm{d}x \int_{y_1}^{y_2} \mathrm{d}y \frac{\partial P(x,y)}{\partial y} = \int_{x_1}^{x_2} \mathrm{d}x \left[P(x, y_1) - P(x, y_2) \right]. \tag{1.318}$$

Adding these together and identifying the right-hand side as the right-hand side of Eq. (1.316) proves Green's theorem for any rectangle.

To prove this for an arbitrary surface S bounded by the contour Γ such as depicted in Fig. 1.17, you fill the surface S with small squares as shown in Fig. 1.17 and apply Green's theorem as just proven to each such small square. The surface integral over

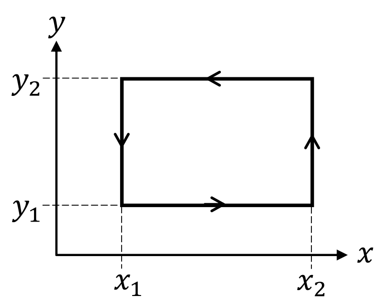

Figure 1.16 Simple rectangle used for proving Green's theorem.

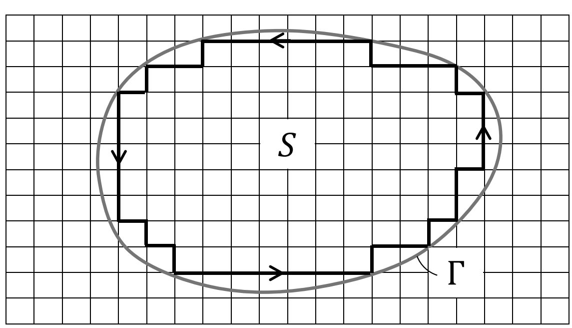

Figure 1.17 Surface S bounded by some contour Γ and filled with small squares made arbitrarily small.

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

any arbitrary S in Eq. (1.316) (the left-hand side) is obtained by adding the surface integrals over all the small squares together. After applying Green's theorem as proven above for each small square, adjacent squares have line integrals on their shared side that cancel such that the only contribution of the line integrals from the sum of small squares is the integral along the heavy jagged line shown in Fig. 1.17. In the limit as the small squares become quite small, the heavy jagged line becomes indistinguishable from the closed contour Γ and the theorem is proven for any S and not just rectangles.

8. Prove the following tensor identity involving the scalar field $\rho(\boldsymbol{r})$ (this identity is needed in Chapter 7 for the development of the differential rules that control how fluid density varies across a meniscus separating two distinct fluids)

$$\nabla \cdot \left[-\frac{1}{2}|\nabla\rho|^2 \boldsymbol{I} + (\nabla\rho)(\nabla\rho) \right] = (\nabla^2\rho)(\nabla\rho), \tag{1.319}$$

where $\boldsymbol{I}$ is again the identity tensor and $|\nabla\rho|^2 = (\nabla\rho)\cdot(\nabla\rho)$. Do this using the method given in Section 1.7. With $\boldsymbol{E} = \nabla\rho$, this is also an identity developed in the proof of the Maxwell stress tensor of Chapter 3.

9. On a surface $\partial\Omega$ that surrounds some region of space Ω and has a normal vector $\boldsymbol{n}$, show that

$$\boldsymbol{n} \times \nabla \times \boldsymbol{E} = \boldsymbol{n} \cdot [\boldsymbol{I}(\nabla \cdot \boldsymbol{E}) - \nabla \boldsymbol{E}] \tag{1.320}$$

for some vector (or tensor) field $\boldsymbol{E}$ distributed throughout Ω and on $\partial\Omega$.

10. With r being radial distance from the origin, demonstrate that

$$\nabla\nabla\left(\frac{1}{r}\right) = -\frac{1}{r^3}\left(\boldsymbol{I} - 3\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}\right), \tag{1.321}$$

where you will need to know that $\nabla r = \hat{\boldsymbol{r}}$, $\boldsymbol{r} = r\hat{\boldsymbol{r}}$ and $\nabla\boldsymbol{r} = \boldsymbol{I}$ (the second-order identity tensor). Note that you do not need to work in spherical coordinates to prove this. This is a needed result once we treat the elastostatic response of a solid.

11. In spherical coordinates, with the position vector defined as $\boldsymbol{r} = r\hat{\boldsymbol{r}}$, demonstrate that

$$\nabla\boldsymbol{r} = \hat{\boldsymbol{r}}\hat{\boldsymbol{r}} + \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\phi}}\hat{\boldsymbol{\phi}} = \boldsymbol{I}. \tag{1.322}$$

12. For the three *fourth-order identity tensors* defined by

$$_4\boldsymbol{I}^{(1)} = \delta_{il}\delta_{jk}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l = \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i \tag{1.323}$$

$$_4\boldsymbol{I}^{(2)} = \delta_{ik}\delta_{jl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l = \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j \tag{1.324}$$

$$_4\boldsymbol{I}^{(3)} = \delta_{ij}\delta_{kl}\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_l = \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_k = \boldsymbol{I}\boldsymbol{I} \tag{1.325}$$

demonstrate that for any second-order tensor $\boldsymbol{A}$

$$_4\boldsymbol{I}^{(1)} : \boldsymbol{A} = \boldsymbol{A} \tag{1.326}$$

$$_4\boldsymbol{I}^{(2)} : \boldsymbol{A} = \boldsymbol{A}^T \tag{1.327}$$

$$_4\boldsymbol{I}^{(3)} : \boldsymbol{A} = \mathrm{tr}\{\boldsymbol{A}\}\,\boldsymbol{I}. \tag{1.328}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

These three fourth-order identity tensors have components that are entirely independent of the coordinates being used and can also be called the fourth-order *isotropic* tensors. Of the 4! possible ways of distributing the indices i, j, k, and l over two Kronecker delta functions, there are only three unique ways as given in Eqs (1.323)–(1.325). So there are three and only three fourth-order isotropic tensors, and we will use them later when we derive the laws of elasticity in an isotropic solid.

13. Derive the fourth-order tensor identity

$$\nabla\nabla\,(\boldsymbol{rr}) = \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i = {}_4\boldsymbol{I}^{(2)} + {}_4\boldsymbol{I}^{(1)}. \tag{1.329}$$

To do so, write each $\nabla = \hat{\boldsymbol{x}}_i\partial/\partial x_i$ and $\boldsymbol{r} = x_i\hat{\boldsymbol{x}}_i$ in Cartesian coordinates with each vector having its own index and use the fact that $\partial x_i/\partial x_j = \delta_{ij}$. Similarly derive the sixth-order tensor identity

$$\begin{aligned}\nabla\nabla\nabla\,(\boldsymbol{rrr}) = {} & \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_j \\ & + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_k + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_i \\ & + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j + \hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_i\end{aligned} \tag{1.330}$$

by again writing each ∇ and $\boldsymbol{r}$ in Cartesian coordinates. Then show that for some third-order tensor ${}_3\boldsymbol{A} = A_{lmn}\hat{\boldsymbol{x}}_l\hat{\boldsymbol{x}}_m\hat{\boldsymbol{x}}_n$ that

$$[\nabla\nabla\nabla\,(\boldsymbol{rrr})] \overset{3}{\vdots}\, {}_3\boldsymbol{A} = \left(A_{ijk} + A_{ikj} + A_{jik} + A_{jki} + A_{kij} + A_{kji}\right)\hat{\boldsymbol{x}}_i\hat{\boldsymbol{x}}_j\hat{\boldsymbol{x}}_k \tag{1.331}$$

$$= \boldsymbol{A} + \boldsymbol{A}^{\overset{T}{132}} + \boldsymbol{A}^{\overset{T}{213}} + \boldsymbol{A}^{\overset{T}{231}} + \boldsymbol{A}^{\overset{T}{312}} + \boldsymbol{A}^{\overset{T}{321}}. \tag{1.332}$$

Such transpose identities are used in the derivation of the Taylor series coefficients for Taylor series of fields (scalars, vectors, or tensors) in three-dimensional space as described in Section 1.8.1.

14. Demonstrate through matrix multiplication that the Cartesian-coordinate rotation matrix for counterclockwise rotations θ_1, θ_2, and θ_3 about the Cartesian axes x_1, x_2, and x_3 is given by

$$R_{ij}(\theta_1, \theta_2, \theta_3) = R_{ik}(\theta_1)R_{kl}(\theta_2)R_{lj}(\theta_3)$$
$$= \begin{bmatrix} \cos\theta_2\cos\theta_3 & \cos\theta_2\sin\theta_3 & -\sin\theta_2 \\ \sin\theta_1\sin\theta_2\cos\theta_3 - \cos\theta_1\sin\theta_3 & \sin\theta_1\sin\theta_2\sin\theta_3 + \cos\theta_1\cos\theta_3 & \sin\theta_1\cos\theta_2 \\ \cos\theta_1\cos\theta_3\sin\theta_2 - \sin\theta_1\sin\theta_3 & \cos\theta_1\sin\theta_2\sin\theta_3 - \sin\theta_1\cos\theta_3 & \cos\theta_1\cos\theta_2 \end{bmatrix}$$

where $R_{ik}(\theta_1)$, $R_{kl}(\theta_2)$, and $R_{lj}(\theta_3)$ are given by Eqs (1.145), (1.152), and (1.153). Then show by direct matrix multiplication that for any rotation of the Cartesian coordinates, the identity tensor in the rotated coordinates is $R_{ik}(\theta_1, \theta_2, \theta_3)R_{jl}(\theta_1, \theta_2, \theta_3)\delta_{kl} = R_{ik}(\theta_1, \theta_2, \theta_3)\left[R_{jk}(\theta_1, \theta_2, \theta_3)\right]^T = \delta_{ij}$. So the second-order identity tensor is an isotropic second-order tensor, that is, a second-order tensor whose components do not change when we make arbitrary changes to the orientation of the axes.

15. For arbitrary orthogonal-curvilinear coordinates having metrical coefficients $h_1(q_1, q_2, q_3)$, $h_2(q_1, q_2, q_3)$, and $h_3(q_1, q_2, q_3)$ as well as unit base vectors $\hat{\boldsymbol{q}}_1$, $\hat{\boldsymbol{q}}_2$, and $\hat{\boldsymbol{q}}_3$, demonstrate that

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

$$\nabla\nabla\,(\boldsymbol{rr}) = \hat{\boldsymbol{q}}_i\hat{\boldsymbol{q}}_j\hat{\boldsymbol{q}}_i\hat{\boldsymbol{q}}_j + \hat{\boldsymbol{q}}_i\hat{\boldsymbol{q}}_j\hat{\boldsymbol{q}}_j\hat{\boldsymbol{q}}_i \tag{1.333}$$

using the ideas developed in Section 1.8.6.

16. Using the well-known definite integral $\int_{-\infty}^{\infty} e^{-u^2}\,\mathrm{d}u = \sqrt{\pi}$ that we will prove in Chapter 10, demonstrate that the representation of the Dirac delta function given by

$$\delta(x - x_1) = \lim_{\sigma\to 0} \frac{e^{-(x-x_1)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}} \tag{1.334}$$

indeed possesses the required property of the Dirac delta that

$$\int_{-\infty}^{\infty} \frac{e^{-(x-x_1)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}\,\mathrm{d}x = 1 \tag{1.335}$$

for any σ. Taking $\sigma \to 0$ is what allows us to obtain the other key property of the Dirac delta that $\int_{-\infty}^{\infty} f(x)\delta(x - x_1)\,\mathrm{d}x = f(x_1)$, which is called the sifting property.

If we now define the Dirac delta to be some arbitrary power n of the above bell-shaped curve

$$\delta(x - x_1) = \lim_{\sigma\to 0} \frac{e^{-n(x-x_1)^2/(2\sigma^2)}}{c_\sigma}, \tag{1.336}$$

show that the normalization constant c_σ that allows this representation to possess the required property $\int_{-\infty}^{\infty} \delta(x - x_1)\,\mathrm{d}x = 1$ is

$$c_\sigma = \sigma\sqrt{\frac{2\pi}{n}}. \tag{1.337}$$

17. Given, say, a second-order tensor field $\boldsymbol{A}(\boldsymbol{r})$ (but this could also be a scalar or vector field) and a point $\boldsymbol{r}_s$ located somewhere within a volumetric region Ω and not on the boundary $\partial\Omega$, demonstrate the 3D gradient-sifting property of the 3D Dirac delta that states

$$\nabla \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r}=\boldsymbol{r}_s} = -\int_\Omega \mathrm{d}\boldsymbol{r}^3\ [\nabla\delta(\boldsymbol{r} - \boldsymbol{r}_s)]\ \boldsymbol{A}(\boldsymbol{r}). \tag{1.338}$$

To do so, you will need to use the fundamental theorem of 3D calculus and the sifting property of the 3D Dirac delta function.

18. Show that if you represent the 3D Dirac delta function using the Gaussian function in the limit as the standard of deviation becomes small, you can represent its gradient as

$$\nabla\delta(\boldsymbol{r} - \boldsymbol{r}_s) = \lim_{\sigma\to 0} -\frac{(\boldsymbol{r} - \boldsymbol{r}_s)}{\sigma^5(2\pi)^{3/2}} \exp\left(\frac{-|\boldsymbol{r} - \boldsymbol{r}_s|^2}{2\sigma^2}\right). \tag{1.339}$$

19. For the integral $I = \int_{-\infty}^{\infty} \mathrm{d}x\,\delta[g(x)]f(x)$, when $g(x) = \sin(\pi x/L)$ and $f(x) = e^{-x/L}$, show that

$$I = \frac{eL}{\pi(e - 1)}. \tag{1.340}$$

https://doi.org/10.1017/9781108951982.003 Published online by Cambridge University Press

HINT: it will prove useful to remember the binomial expansion $(1-u)^{-1} = \sum_{n=0}^{\infty} u^n$ for $|u| < 1$.

20. Demonstrate that for $a > 0$,

$$\delta(x^2 - a^2) = \frac{1}{2a}\left[\delta(x+a) + \delta(x-a)\right]. \tag{1.341}$$

21. For a symmetric second-order tensor given by $\boldsymbol{S} = \boldsymbol{ab} + \boldsymbol{ba}$ and a so called anti-symmetric second-order tensor given by $\boldsymbol{A} = \boldsymbol{cd} - \boldsymbol{dc}$, where $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, and $\boldsymbol{d}$ are vectors, show that $\boldsymbol{S} : \boldsymbol{A} = 0$.