1

Some mathematical essentials

1.1 Scalars, vectors, and Cartesian tensors

Geometry is a vital ingredient in the description of continuum problems. Our treatment will focus on the mathematically simplest representation for this subject. Although curvilinear coordinates can be more natural, they introduce complications that go beyond the scope of this book. The initial part of our treatment will parallel the Cartesian approach of Mase and Mase (1990) rather than the curvilinear approach of Narasimhan (1993) and Fung (1965). Hence, we will adhere to a Cartesian description of problems and be spared the need to distinguish between covariant and contravariant notation. Moreover, we will generally employ second-rank tensors which are matrices that possess some very special and important (coordinate) transformation properties.

We will distinguish between three classes of objects: namely, scalars, vectors, and tensors. In reality, all quantities may be regarded as tensors of a specific rank. *Scalar* (nonconstant) quantities, such as density and temperature, are *zero rank* or *order* tensors, while *vector* quantities (which have an associated direction, such as velocity) are *first-rank* tensors. *Second-rank tensors*, such as the stress tensor, are a special case of square matrices. We will usually denote vector quantities by bold-face lower-case letters, while second-rank tensors will be denoted by bold-face upper-case letters.

To simplify our geometrical description of problems, we will employ an indicial notation. In lieu of x, y, and z orthogonal axes, we will employ x_1, x_2, and x_3. Similarly, we will denote by $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, and $\hat{\mathbf{e}}_3$ *unit*-vectors in the directions of x_1, x_2, and x_3. The indicial notation implies that any repeated index is implicitly summed, generally from 1 through 3. This is the *Einstein summation convention*. It is sufficient to denote a vector $\mathbf{v}$ by its three components (v_1, v_2, v_3). We conform with the tradition that vector quantities are shown in bold face. For more detail on these issues, the reader is advised to consult mathematical texts such as those by Boas

(2006), Arfken and Weber (2005), Mathews and Walker (1970), or Schutz (1980) ranging from elementary to advanced treatments. We note that $\mathbf{v}$ can be represented vectorially then as

$$\mathbf{v} \equiv \sum_{i=1}^{3} v_i \, \hat{\mathbf{e}}_i = v_i \, \hat{\mathbf{e}}_i . \tag{1.1}$$

Accordingly, we see that

$$T_{ij} \, v_i \, \hat{\mathbf{e}}_j = \sum_{i=1,\, j=1}^{3} T_{ij} v_i \hat{\mathbf{e}}_j . \tag{1.2}$$

Moreover, vectors and matrices have certain of the properties of a group: (a) the sum of two vectors is a vector; (b) the product of a vector by a scalar is a vector, etc. In addition, we define an inner (scalar) product or dot product according to the usual physicist's convention

$$\mathbf{u} \cdot \mathbf{v} \equiv u_i \, v_i . \tag{1.3}$$

It is important to note that mathematicians sometimes employ a slightly different notation for the inner product, namely

$$(\mathbf{u}, \mathbf{v}) = \mathbf{u}^T \mathbf{v} = \mathbf{v}^T \mathbf{u} = u_i \, v_i , \tag{1.4}$$

where they assume that $\mathbf{u}$ and $\mathbf{v}$ are column vectors. The latter formalism requires particular care since transposed quantities often appear. In order to maintain transparency in all of our derivations, we will employ primarily the indicial notation. Moreover, if we define u and v to be the lengths of $\mathbf{u}$ and $\mathbf{v}$, respectively, according to

$$u \equiv \sqrt{u_i \, u_i} = |\mathbf{u}| \, ; \qquad v \equiv \sqrt{v_i \, v_i} = |\mathbf{v}| \, , \tag{1.5}$$

we can identify an angle θ between $\mathbf{u}$ and $\mathbf{v}$ which we define according to

$$\mathbf{u} \cdot \mathbf{v} \equiv u \, v \cos \theta . \tag{1.6}$$

We now introduce the Kronecker delta δ_{ij} and the Levi-Civita permutation symbol ϵ_{ijk} owing to their utility in tensor calculations.

We define the Kronecker delta according to

$$\delta_{ij} \equiv \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} . \tag{1.7}$$

It follows that the Kronecker delta is the realization of the identity matrix. It follows then that

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij} , \tag{1.8}$$

and that

$$\delta_{ii} = 3. \tag{1.9}$$

(This is equivalent to saying that the *trace*, i.e. the sum of the diagonal elements, of the identity matrix is 3.) An important consequence of equation (1.8) is that

$$\delta_{ij}\,\hat{\mathbf{e}}_j = \hat{\mathbf{e}}_i. \tag{1.10}$$

We can employ these definitions to derive the general scalar product relation (1.3) using the special case (1.8). In particular, it follows that

$$\mathbf{u}\cdot\mathbf{v} = u_i\,\hat{\mathbf{e}}_i\cdot v_j\,\hat{\mathbf{e}}_j = u_i\,v_j\,\hat{\mathbf{e}}_i\cdot\hat{\mathbf{e}}_j = u_i\,v_j\,\delta_{ij} = u_i\,v_i. \tag{1.11}$$

In order to introduce the vector (cross) product, we introduce the Levi-Civita or permutation symbol ϵ_{ijk} according to

$$\epsilon_{ijk} \equiv \begin{cases} 1, & \text{if } i\ j\ k \text{ are an even permutation of } 1\,2\,3 \\ -1, & \text{if } i\ j\ k \text{ are an odd permutation of } 1\,2\,3 \\ 0, & \text{if any two of } i,\ j,\ k \text{ are the same} \end{cases}. \tag{1.12}$$

We note that ϵ_{ijk} changes sign if any two of its indices are interchanged. For example, if the 1 and 3 are interchanged, then the sequence 1 2 3 becomes 3 2 1. Accordingly, we *define* the cross product $\mathbf{u}\times\mathbf{v}$ according to its ith component, namely

$$(\mathbf{u}\times\mathbf{v})_i \equiv \epsilon_{ijk}\,u_j\,v_k, \tag{1.13}$$

or, equivalently,

$$\mathbf{u}\times\mathbf{v} = (\mathbf{u}\times\mathbf{v})_i\,\hat{\mathbf{e}}_i = \epsilon_{ijk}\,\hat{\mathbf{e}}_i\,u_j\,v_k = -\,(\mathbf{v}\times\mathbf{u}). \tag{1.14}$$

By inspection, it is observed that this kind of structure is closely connected to the definition of the determinant of a 3×3 matrix. This follows directly when we write the scalar triple product

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \epsilon_{ijk}\,u_i\,v_j\,w_k, \tag{1.15}$$

which, by virtue of the cyclic permutivity of the Levi-Civita symbol demonstrates

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \mathbf{v}\cdot(\mathbf{w}\times\mathbf{u}) = \mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}). \tag{1.16}$$

The right side of equation (1.15) is the determinant of a matrix whose rows correspond to $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, respectively. It is useful to note that the scalar triple product can be employed to establish whether a Cartesian coordinate system is right or left handed, i.e. the product is $+1$ or -1.

It is useful, as well, to consider the vector triple cross product

$$\begin{aligned}\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) &= \mathbf{u}\times\epsilon_{ijk}\,\hat{\mathbf{e}}_i\,v_j\,w_k = \epsilon_{lmi}\,\hat{\mathbf{e}}_l\,u_m\,\epsilon_{ijk}\,v_j\,w_k \\ &= \left(\epsilon_{ilm}\,\epsilon_{ijk}\right)\hat{\mathbf{e}}_l\,u_m\,v_j\,w_k.\end{aligned} \tag{1.17}$$

It is necessary to deal first with the $\epsilon_{ilm}\epsilon_{ijk}$ term. Observe, as we sum over the i index, that contributions can emerge only if $l \neq m$ and $j \neq k$. If these conditions both hold, then we get a contribution of 1 if $l = j$ and $m = k$ and a contribution of -1 if $l = k$ and $m = j$. Hence, it follows that

$$\epsilon_{ilm}\,\epsilon_{ijk} = \delta_{lj}\,\delta_{mk} - \delta_{lk}\,\delta_{mj}. \tag{1.18}$$

Returning to equation (1.17), we obtain that

$$\begin{aligned} \mathbf{u} \times (\mathbf{v} \times \mathbf{w}) &= \left(\delta_{lj}\,\delta_{mk} - \delta_{lk}\,\delta_{mj}\right) \hat{\mathbf{e}}_l\, u_m\, v_j\, w_k \\ &= \hat{\mathbf{e}}_l\, v_l\, u_m\, w_m - \hat{\mathbf{e}}_l\, w_l\, u_m\, v_m \\ &= \mathbf{v}\,(\mathbf{u} \cdot \mathbf{w}) - \mathbf{w}\,(\mathbf{u} \cdot \mathbf{v}), \end{aligned} \tag{1.19}$$

thereby reproducing a familiar, albeit otherwise cumbersome to derive, algebraic identity. Finally, if we replace the role of $\mathbf{u}$ in the triple scalar product (1.15) by $\mathbf{v} \times \mathbf{w}$, it immediately follows that

$$\begin{aligned} (\mathbf{v} \times \mathbf{w}) \cdot (\mathbf{v} \times \mathbf{w}) = |\mathbf{v} \times \mathbf{w}|^2 &= \epsilon_{ijk}\, v_j\, w_k \epsilon_{ilm}\, v_l\, w_m \\ &= \left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right) v_j\, w_k\, v_l\, w_m. \end{aligned} \tag{1.20}$$

Finally, this can be written

$$|\mathbf{v} \times \mathbf{w}|^2 = v^2 w^2 - (\mathbf{v} \cdot \mathbf{w})^2 = v^2\, w^2\, \sin^2\theta, \tag{1.21}$$

where we have made use of the definition (1.6). Indeed, it is possible to demonstrate the validity of many other vector identities by employing the Levi-Civita and Kronecker symbols. This is especially true with respect to *derivative* operators. We define ∂_i according to

$$\partial_i \equiv \frac{\partial}{\partial x_i}. \tag{1.22}$$

Another notational shortcut that is commonly used is to employ a subscript of ", i" to denote a derivative with respect to x_i; importantly, a comma "," is employed to designate differentiation together with the subscript. Hence, if f is a scalar function of $\mathbf{x}$, we write

$$\frac{\partial f}{\partial x_i} = \partial_i f = f_{,i}; \tag{1.23}$$

but if $\mathbf{g}$ is a vector function of $\mathbf{x}$, we write

$$\frac{\partial g_i}{\partial x_j} = \partial_j\, g_i = g_{i,j}. \tag{1.24}$$

Higher derivatives may be expressed using this shorthand as well, e.g.

$$\frac{\partial^2 g_i}{\partial x_j\, \partial x_k} = g_{i,jk}. \tag{1.25}$$

Cambridge University Press
978-0-521-56289-8 - Continuum Mechanics in the Earth Sciences
William I. Newman
Excerpt
More information

Then, the usual gradient, divergence, and curl operators become

$$\nabla = \hat{\mathbf{e}}_i \, \partial_i ; \tag{1.26}$$

$$\nabla \cdot \mathbf{u} = \partial_i \, u_i ; \tag{1.27}$$

and

$$\nabla \times \mathbf{u} = \epsilon_{ijk} \, \hat{\mathbf{e}}_i \, \partial_j \, u_k , \tag{1.28}$$

where $\mathbf{u}$ is a vector and a function of the position vector $\mathbf{x}$. With only a modest degree of additional effort (noting that ∂_i commutes with ∂_j, an added benefit of Cartesian geometry, but not with any *function* of $\mathbf{x}$), it is now relatively simple to derive all vector identities in Cartesian coordinates. Before proceeding further, it deserves mention that the usual theorems of Gauss, Green, and Stokes also hold for tensor quantities albeit in a slightly more complicated form than for vector ones. The essential point here is that as one converts from a volume integral to a surface integral and to a line integral, appropriate differential operators are introduced. In a Cartesian representation, these rarely cause any problems and the rules discussed above apply.

One final notation issue needs to be addressed at this time, the tensor (outer) product of two vectors, the so-called *dyad*

$$\mathbf{u}\,\mathbf{v} = u_i \, \hat{\mathbf{e}}_i \, v_j \, \hat{\mathbf{e}}_j = u_i \, v_j \, \hat{\mathbf{e}}_i \, \hat{\mathbf{e}}_j . \tag{1.29}$$

(Although some engineering texts sometimes introduce the $\otimes$ symbol between the two vector quantities, mathematics texts often employ the $\otimes$ symbol to denote an "antisymmetric" form, i.e. $\mathbf{u} \otimes \mathbf{v} = \mathbf{uv} - \mathbf{vu}$. Other mathematics texts employ the "wedge" $\mathbf{u} \wedge \mathbf{v}$ for this purpose. The absence of an intervening dot between the unit vectors is significant: there is no inner product implied.) Operations on dyads follow the usual rules for the relevant vector components; e.g.

$$\mathbf{u}\,\mathbf{v} \cdot \mathbf{p}\,\mathbf{q} = \mathbf{u}\,(\mathbf{v} \cdot \mathbf{p})\,\mathbf{q} = (\mathbf{v} \cdot \mathbf{p})\,\mathbf{u}\,\mathbf{q} , \tag{1.30}$$

where we note that the scalar product $(\mathbf{v} \cdot \mathbf{p})$ should be regarded solely as a scalar quantity. Dyads are particularly useful in the decomposition or diagonalization of matrices. Note, also, that although a dyad has nine components, only six independent quantities are involved (and these six quantities can be calculated up to a multiplicative constant). It is possible in similar fashion to construct triadic, tetradic, and higher rank tensors.

Thus far, our treatment of geometry and some of the underpinnings of scalars, vectors and Cartesian tensors has been abstract. The power of these methods is greatly enhanced when we employ our geometric intuition in solving problems. (This philosophy is also at the heart of our adherence to Cartesian coordinates in our treatment. Once we have derived the fundamental equations of continuum mechanics in Cartesian form, it is relatively simple to convert them to other curvilinear coordinate systems, such as cylindrical or spherical. What we gain in the process is geometrical simplicity.)

As an illustration of the utility of coupling geometric intuition with the formalism, consider the classic problem of establishing the bond angles in a methane or CH_4 molecule. Visualize the carbon atom at the origin of our coordinate system and let us assume that one of the hydrogen atoms is at a distance v from the carbon atom along the z-axis. We will designate its position by the vector $\mathbf{v}^{(0)}$. The three remaining hydrogen atoms make an angle θ, to be determined, with respect to the hydrogen atom on the z-axis, and we will call their vectors $\mathbf{v}^{(i)}$ for $i = 1, 2, 3$. The "center of mass" of the hydrogen atoms is $\sum_{i=0}^{3} \mathbf{v}^{(i)} = \mathbf{0}$, i.e. at the origin where the carbon atom is situated. We take the dot product of $\mathbf{v}^{(0)}$ with this latter quantity to find that $v^2 + 3\,v^2 \cos\theta = 0$ leaving $\cos\theta = -1/3$ or $\theta \approx 109.471\,220\,634°$. Drawing a picture is *always* a good idea. Let us consider a more complex example.

As a novel illustration of the use of dyads, consider the "corner cube" reflector frequently employed on highways and elsewhere owing to their remarkable ability to reflect back to the source light incident on the device from *any* angle.

A corner reflector consists of three mutually perpendicular, intersecting flat surfaces (see figure 1.1). In particular, assume that the corner's faces can be described by normal unit vectors $\hat{\mathbf{n}}_i$, $i = 1, 2, 3$ as shown in the accompanying figure. If a light ray with direction $\hat{\mathbf{r}}$ impinges on face #1, then its projection along the $\hat{\mathbf{n}}_1$ direction is reversed while its projection on the $\hat{\mathbf{n}}_i$ for $i = 2, 3$ remains unchanged. Thus, the new vector $\hat{\mathbf{r}}'$ after the first reflection is given by

$$\hat{\mathbf{r}}' = \left(\hat{\mathbf{n}}_2\,\hat{\mathbf{n}}_2 + \hat{\mathbf{n}}_3\,\hat{\mathbf{n}}_3 - \hat{\mathbf{n}}_1\,\hat{\mathbf{n}}_1\right) \cdot \hat{\mathbf{r}} = \left(\mathbf{I} - 2\,\hat{\mathbf{n}}_1\,\hat{\mathbf{n}}_1\right) \cdot \hat{\mathbf{r}}, \qquad (1.31)$$

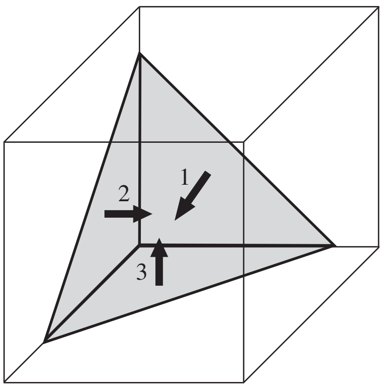

Figure 1.1 Geometry of a corner cube retroreflector.

where

$$\mathbf{I} = \hat{\mathbf{n}}_1\,\hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2\,\hat{\mathbf{n}}_2 + \hat{\mathbf{n}}_3\,\hat{\mathbf{n}}_3 \tag{1.32}$$

defines the identity operator. We now allow for a reflection on face number 2, yielding $\hat{\mathbf{r}}''$. Taking all three reflections in turn, we obtain for the final ray vector

$$\hat{\mathbf{r}}''' = \left(\mathbf{I} - 2\,\hat{\mathbf{n}}_3\,\hat{\mathbf{n}}_3\right)\cdot\left(\mathbf{I} - 2\,\hat{\mathbf{n}}_2\,\hat{\mathbf{n}}_2\right)\cdot\left(\mathbf{I} - 2\,\hat{\mathbf{n}}_1\,\hat{\mathbf{n}}_1\right)\cdot\hat{\mathbf{r}} = -\hat{\mathbf{r}}, \tag{1.33}$$

after a modest amount of algebra where the scalar products of the form $\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2$, etc., are observed to vanish. Hence, we see how the utilization of these geometrical constructs can dramatically simplify the solution of practical problems.[1]

The notation employed thus far has mixed symbolic and indicial conventions; the proper treatment requires a covariant–contravariant formulation which would also accommodate curvilinear coordinates. Engineering texts generally employ notation that preserves vectorial quantities using bold-face characters, e.g. $\mathbf{n}$, while physics texts such as Landau *et al.* (1986) and Landau and Lifshitz (1987) generally employ indicial notation, e.g. n_i. When referring to tensors (and to matrices), the symmetries possessed with respect to the indices can be particularly important. For example, if $A_{ij} = -A_{ji}$, we say that the second-rank tensor $\mathbf{A}$ is anti-symmetric; similarly, if $A_{ij} = A_{ji}$, we say that $\mathbf{A}$ is symmetric. To maintain a (subtle) distinction between matrices and tensors, we shall denote the associated matrix by the symbol $\mathcal{A}$. We turn now to some of the essential properties of matrices and determinants.

1.2 Matrices and determinants

A *matrix*, unlike a tensor which properly defined also preserves coordinate transformation properties, is an ordered rectangular array of elements enclosed by brackets. The reader may wish to review a good linear algebra text before continuing further. An encyclopedic source of information concerning matrices is the two-volume text by Gantmakher (1959). An M by N matrix (written $M \times N$) can be expressed

$$\mathcal{A} = \left[A_{ij}\right] = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \dots & A_{MN} \end{pmatrix}. \tag{1.34}$$

[1] An especially elegant way of demonstrating the reversal of the direction of the incident beam is to adopt temporarily a rotated coordinate system whose axes are orthogonal to each of the three cube faces. In this system, each of the components of the ray vector undergoes a reversal as the light ray encounters each cube face, respectively. Since all three components of the initial vector are reversed after the three reflections, the outcome of this encounter with a corner cube is the reversal of the light ray.

A *zero* or *null* matrix has all elements equal to zero. A *diagonal* matrix is a square matrix whose elements which are not on the *principal diagonal* vanish; the *unit* or *identity* matrix is a diagonal matrix whose diagonal elements are unit values. By interchanging the rows and columns of an $M \times N$ matrix $\mathcal{A}$, we form its $N \times M$ transpose $\mathcal{A}^T$. It follows then that a symmetric matrix is a square matrix $\mathcal{A} = \mathcal{A}^T$. Any matrix can be expressed as the sum of a symmetric and an antisymmetric matrix, respectively, namely

$$\mathcal{A} = \frac{\mathcal{A} + \mathcal{A}^T}{2} + \frac{\mathcal{A} - \mathcal{A}^T}{2}. \tag{1.35}$$

For complex-valued matrices, we sometimes denote by the superscript $\dagger$ or H the complex conjugate transpose or *Hermitian* operator. Thus, we define $\mathcal{A}^\dagger \equiv \mathcal{A}^H \equiv \mathcal{A}^{T\star}$ where a $\star$ is used to indicate the complex conjugate. We say that a matrix $\mathcal{A}$ is Hermitian if it is identical to its conjugate transpose. Such properties sometimes emerge in the manipulation of matrices, but rarely so in continuum mechanics.

Matrix addition is commutative, i.e. $\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$, and associative, i.e. $\mathcal{A} + (\mathcal{B} + \mathcal{C}) = (\mathcal{A} + \mathcal{B}) + \mathcal{C}$. Multiplication of a matrix $\mathcal{A}$ by a scalar λ gives rise to a new matrix $\lambda\,\mathcal{A}$. A matrix product $\mathcal{C} = \mathcal{A}\,\mathcal{B}$ can be defined in a manner reminiscent of an inner product, that is

$$C_{ij} = A_{ik}\, B_{kj}, \tag{1.36}$$

where summation over the index k takes place over the admissible range – note that the matrices $\mathcal{A}$ and $\mathcal{B}$ must be compatible in size. Observe further that matrix multiplication is *not* commutative, i.e. $\mathcal{A}\,\mathcal{B} \neq \mathcal{B}\,\mathcal{A}$.

It is often useful to define a *quadratic form* from a matrix $\mathcal{A}$ or its equivalent second rank tensor $\mathbf{A}$, namely $\mathbf{x}^H\,\mathcal{A}\,\mathbf{x}$. We say that a matrix is positive definite if $\mathbf{x}^H\,\mathcal{A}\,\mathbf{x} > 0$ for all $\mathbf{x}$ and that it is positive semi-definite if $\mathbf{x}^H\,\mathcal{A}\,\mathbf{x} \geq 0$. (This property of matrices is of fundamental importance in continuum mechanics and in stability theory.) The sum of the diagonal elements of a matrix is its trace, i.e. $\operatorname{tr}\mathcal{A} \equiv A_{ii}$. For convenience, we will define the determinant of a matrix using the Levi-Civita symbol, extending the definition (1.12) of the symbol to an arbitrary number N of indices (corresponding to an $N \times N$ matrix). In particular, we define

$$\epsilon_{i_1 i_2 \ldots i_N} \equiv \begin{cases} 1, & \text{if } i_1 i_2 \ldots i_N \text{ are an even permutation of } 1, 2, \ldots N \\ -1, & \text{if } i_1 i_2 \ldots i_N \text{ are an odd permutation of } 1, 2, \ldots N. \\ 0, & \text{if any two of } i_1 i_2 \ldots i_N \text{ are the same} \end{cases} \tag{1.37}$$

The issue of even and odd permutations can be understood this way: if we interchange any two indices of the Levi-Civita symbol, then it changes sign, alternating between $+1$ and -1. Accordingly, however, we note that if any two indices are the same, then the Levi-Civita symbol must vanish. Then, we define $\det \mathcal{A}$ according to

$$\det \mathcal{A} \equiv \epsilon_{i_1 i_2 \ldots i_N} A_{1i_1} A_{2i_2} \ldots A_{Ni_N}. \tag{1.38}$$

It is easy to show that this reduces to the familiar definition for the determinant of a 3×3 matrix. Moreover, it is easy to show that all of the familiar properties of a determinant are preserved through this definition – particularly, if we regard each row (or column) of $\mathcal{A}$ as a vector, then the determinant vanishes if any row (or column) can be expressed as a linear combination of the other rows (or columns). Finally, these rules can be employed to develop the usual rules for the evaluation of a determinant by cofactors, etc. The inverse of a matrix $\mathcal{A}^{-1}$, *if it exists*, is a matrix defined such that $\mathcal{A}^{-1}\mathcal{A} = \mathcal{A}\mathcal{A}^{-1} = \mathcal{I}$, the identity matrix. We observe that the transpose or inverse of a product reverses the usual order of terms, that is

$$(\mathcal{A}\mathcal{B})^T = \mathcal{B}^T \mathcal{A}^T, \tag{1.39}$$

and

$$(\mathcal{A}\mathcal{B})^{-1} = \mathcal{B}^{-1} \mathcal{A}^{-1}. \tag{1.40}$$

Also, it is important to note that the determinant of a product is equal to the product of the determinants, namely

$$\det(\mathcal{A}\mathcal{B}) = (\det \mathcal{A})(\det \mathcal{B}). \tag{1.41}$$

Although harder to prove, this is demonstrable using the Levi-Civita expression for the determinant of a matrix.

1.3 Transformations of Cartesian tensors

It is possible to convert (or rotate from) one coordinate system to another via a "transformation" – the existence of this transformation for tensors but not in general for matrices is what makes tensors a special case of matrices. Suppose we have one set of coordinate axes defined by unit vectors $\hat{\mathbf{e}}_i$ and wish to transform to another set of coordinate axes defined by unit vectors $\hat{\mathbf{e}}'_i$. Then, we can write (since we are dealing with a linear superposition)

$$\hat{\mathbf{e}}'_i = A_{ij}\,\hat{\mathbf{e}}_j, \tag{1.42}$$

where it follows that

$$A_{ij} = \hat{\mathbf{e}}'_i \cdot \hat{\mathbf{e}}_j. \tag{1.43}$$

Thus, we see that the coefficients of the transformation matrix are just the "direction cosines" defining the angles between the old and new coordinate systems. Expression (1.42) will be employed universally in transforming vectors and, later, tensors from one Cartesian coordinate system to another. Similarly, we can write for the inverse transform

$$\hat{\mathbf{e}}_i = A^{-1}_{ij}\,\hat{\mathbf{e}}'_j, \tag{1.44}$$

where

$$A^{-1}_{ij} = \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}'_j. \tag{1.45}$$

By inspection, we observe that

$$\mathcal{A}^{-1} = \mathcal{A}^T. \tag{1.46}$$

The action of $\mathcal{A}$ on a vector can be regarded, alternately, as a rotation from one coordinate system to another (i.e. from the unprimed to the primed), or as a physical rotation of the vector in a manner specified by the reorientation of the coordinate axes.

The matrix $\mathcal{A}$ has some notable properties and, in particular, is referred to as *unitary* or length preserving. This can be shown by observing that, if

$$\mathbf{u}' \equiv \mathbf{A} \cdot \mathbf{u} = \mathcal{A}\,\mathbf{u}, \tag{1.47}$$

or, using indicial notation,

$$u'_i = A_{ij}\,u_j, \tag{1.48}$$

then

$$u'^2 = \mathbf{u}'^T \cdot \mathbf{u}' = \mathbf{u}^T \cdot \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{u} = \mathbf{u}^T \cdot \mathbf{u} = u^2, \tag{1.49}$$

for *any* $\mathbf{u}$. Another important outcome of equations (1.42) and (1.48) emerges due to

$$\mathbf{u}' = u'_i\,\hat{\mathbf{e}}'_i = A_{ij}\,u_j\,A_{ik}\,\hat{\mathbf{e}}_k = A^T_{ji}\,A_{ik}\,u_j\,\hat{\mathbf{e}}_k = \delta_{jk}\,u_j\,\hat{\mathbf{e}}_k = \mathbf{u}, \tag{1.50}$$

thereby showing that the two representations that we have for our original vector $\mathbf{u}$ are identical. The transformation equation (1.48) allows us to convert almost effortlessly from one coordinate system to the other without changing any physical quantities.

If we regard each column of $\mathcal{A}$ as being a vector, say $\mathbf{u}_{(i)}$, $i = 1, 2, 3$, a little manipulation produces (since the inverse and transpose of $\mathcal{A}$ are identical) the result

$$\mathbf{u}_{(i)} \cdot \mathbf{u}_{(j)} = \delta_{ij}. \tag{1.51}$$

Thus, these three vectors are both orthogonal and of unit length. Similarly, one can regard each row of $\mathcal{A}$ as being a vector, say $\mathbf{v}_{(i)}$, and similarly show that these three vectors are also orthogonal and of unit length. Note that we used inner-product notation in the context of a tensor, i.e., we considered the vector $\mathbf{A} \cdot \mathbf{u}$ which is