1
Vectors and Tensors

The mechanics of solids is a story told in the language of vectors and tensors. These abstract mathematical objects provide the basic building blocks of our analysis of the behavior of solid bodies as they deform and resist force. Anyone who stands poised to undertake the study of structural mechanics has undoubtedly encountered vectors at some point. However, in an effort to establish a least common denominator among readers, we shall do a quick review of vectors and how they operate. This review serves the auxiliary purpose of setting up some of the notational conventions that will be used throughout the book.

Our study of mechanics will naturally lead us to the concept of the tensor, which is a subject that may be less familiar (possibly completely unknown) to the reader who has the expected background knowledge in elementary mechanics of materials. We shall build the idea of the tensor from the ground up in this chapter with the intent of developing a facility for tensor operations equal to the facility that most readers will already have for vector operations. In this book we shall be content to stick with a Cartesian view of tensors in rectangular coordinate systems. General tensor analysis is a mathematical subject with great beauty and deep significance. However, the novice can be blinded by its beauty to the point of missing the simple physical principles that are the true subject of mechanics. So we shall cling to the simplest possible rendition of the story that still respects the tensorial nature of solid mechanics.

Mathematics is the natural language of mechanics. This chapter presents a fairly brief treatment of the mathematics we need to start our exploration of solid mechanics. In particular, it covers some basic algebra and calculus of vectors and tensors. Plenty more math awaits us in our study of structural me-
Fundamentals of Structural Mechanics

This chapter lays the foundation of the mathematical notation that we will use throughout the book. As such, it is both a starting place and a refuge to regain one's footing when the going gets tough.

The Geometry of Three-dimensional Space

We live in three-dimensional space, and all physical objects that we are familiar with have a three-dimensional nature to their geometry. In addition to solid bodies, there are basically three primitive geometric objects in three-dimensional space: the point, the curve, and the surface. Figure 1 illustrates these objects by taking a slice through the three-dimensional solid body \( \mathcal{B} \) (a cube, in this case). A point describes position in space, and has no dimension or size. The point \( \mathcal{P} \) in the figure is an example. The most convenient way to describe the location of a point is with a coordinate system like the one shown in the figure. A coordinate system has an origin \( \mathcal{O} \) (a point whose location we understand in a deeper sense than any other point in space) and a set of three coordinate directions that we use to measure distance. Here we shall confine our attention to Cartesian coordinates, wherein the coordinate directions are mutually perpendicular. The location of a point is then given by its coordinates \( \mathbf{x} = (x_1, x_2, x_3) \). A point has a location independent of any particular coordinate system. The coordinate system is generally introduced for the convenience of description or numerical computation.

A curve is a one-dimensional geometric object whose size is characterized by its arc length. In a sense, a curve can be viewed as a sequence of points. A curve has some other interesting properties. At each point along a curve, the curve seems to be heading in a certain direction. Thus, a curve has an orientation in space that can be characterized at any point along the curve by the line tangent to the curve at that point. Another property of a curve is the rate at which this orientation changes as we move along the curve. A straight line is

![Figure 1](image_url)
a curve whose orientation never changes. The curve C exemplifies the geometric notion of curves in space.

A surface is a two-dimensional geometric object whose size is characterized by its surface area. In a certain sense, a surface can be viewed as a family of curves. For example, the collection of lines parallel and perpendicular to the curve C constitute a family of curves that characterize the surface S. A surface can also be viewed as a collection of points. Like a curve, a surface also has properties related to its orientation and the rate of change of this orientation as we move to adjacent points on the surface. The orientation of a surface is completely characterized by the single line that is perpendicular to the tangent lines of all curves that pass through a particular point. This line is called the normal direction to the surface at the point. A flat surface is usually called a plane, and is a surface whose orientation is constant.

A three-dimensional solid body is a collection of points. At each point, we ascribe some physical properties (e.g., mass density, elasticity, and heat capacity) to the body. The mathematical laws that describe how these physical properties affect the interaction of the body with the forces of nature summarize our understanding of the behavior of that body. The heart of the concept of continuum mechanics is that the body is continuous, that is, there are no finite gaps between points. Clearly, this idealization is at odds with particle physics, but, in the main, it leads to a workable and useful model of how solids behave. The primary purpose of hanging our whole theory on the concept of the continuum is that it allows us to do calculus without worrying about the details of material constitution as we pass to infinitesimal limits. We will sometimes find it useful to think of a solid body as a collection of lines, or a collection of surfaces, since each of these geometric concepts builds from the notion of a point in space.

Vectors

A vector is a directed line segment and provides one of the most useful geometric constructs in mechanics. A vector can be used for a variety of purposes. For example, in Fig. 2 the vector v records the position of point b relative to point a. We often refer to such a vector as a position vector, particularly when a is the origin of coordinates. Close relatives of the position vector are displacement (the difference between the position vectors of some point at different times), velocity (the rate of change of displacement), and acceleration (the rate of change of velocity). The other common use of the notion of a vector, to which we shall appeal in this book, is the concept of force. We generally think

![Figure 2](image-url)
Fundamentals of Structural Mechanics

of force as an action that has a magnitude and a direction. Likewise, displace­ments are completely characterized by their magnitude and direction. Because a vector possesses only the properties of magnitude (length of the line) and di­rection (orientation of the line in space), it is perfectly suited to the mathemati­cal modeling of things like forces and displacements. Vectors have many other uses, but these two are the most important in the present context.

Graphically, we represent a vector as an arrow. The shaft of the arrow gives the orientation and the head of the arrow distinguishes the direction of the vec­tor from the two possibilities inherent in the line segment that describes the shaft (i.e., line segments \(ab\) and \(ba\) in Fig. 2 are both oriented the same way in space). The length, or magnitude, of a vector \(\mathbf{v}\) is represented graphically by the length of the shaft of the arrow and will be denoted symbolically as \(\|\mathbf{v}\|\) throughout the book.

The magnitude and direction of a vector do not depend upon any coordinate system. However, for computation it is most convenient to describe a vector in relation to a coordinate system. For that purpose, we endow our coordinate system with unit base vectors \(\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}\) pointing in the direction of the coor­dinate axes. The base vectors are geometric primitives that are introduced purely for the purpose of establishing the notion of direction. Like the origin of coordinates, we view the base vectors as vectors that we understand more deeply and intuitively than any other vector in space. Basically, we assume that we know what it means to be pointing in the \(\mathbf{e}_1\) direction, for example. Any collection of three vectors that point in different directions makes a suitable ba­sis (in the language of linear algebra we would say that three such vectors span three-dimensional space). Because we have introduced the notion of base vec­tors for convenience, we shall adopt the most convenient choice. Throughout this book, we will generally employ orthogonal unit vectors in conjunction with a Cartesian coordinate system.

Any vector can be described in terms of its components relative to a set of base vectors. A vector \(\mathbf{v}\) can be written in terms of base vectors \(\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}\) as

\[
\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3
\]

where \(v_1, v_2,\) and \(v_3\) are called the components of the vector relative to the ba­sis. The component \(v_1\) measures how far the vector extends in the \(\mathbf{e}_1\) direction, as shown in Fig. 3. A component of a vector is a scalar.

Vector operations. An abstract mathematical construct is not really useful until you know how to operate with it. The most elementary operations in mathematics are addition and multiplication. We know how to do these opera­tions for scalars; we must establish some corresponding operations for vectors.

Vector addition is accomplished with the head-to-tail rule or parallelogram rule. The sum of two vectors \(\mathbf{u}\) and \(\mathbf{v}\), which we denote \(\mathbf{u} + \mathbf{v}\), is the vector con­necting the tail of \(\mathbf{u}\) with the head of \(\mathbf{v}\) when the tail of \(\mathbf{v}\) lies at the head of \(\mathbf{u}\), as shown in Fig. 4. If the vectors \(\mathbf{u}\) and \(\mathbf{v}\) are replicated to form the sides of a
parallelogram $abcd$, then $u + v$ is the diagonal $ac$ of the parallelogram. Subtraction of vectors can be accomplished by introducing the negative of a vector, $-v$ (segment $bf$ in Fig. 4), as a vector with the same magnitude that points in exactly the opposite direction of $v$. Then, $u - v$ is simply realized as $u + (-v)$. If we construct another parallelogram $abfe$, then $u - v$ is the diagonal $af$. It is evident from the figure that segment $af$ is identical in length and direction to segment $db$. A vector can be added to another vector, but a vector and a scalar cannot be added (the well-worn analogy of the impossibility of adding apples and oranges applies here).

We can multiply a vector $v$ by a scalar $\alpha$ to get a vector $\alpha v$ having the same direction but a length equal to the original length $\| v \|$ multiplied by $\alpha$. If the scalar $\alpha$ has a negative value, then the sense of the vector is reversed (i.e., it puts the arrow head on the other end). With these definitions, we can make sense of Eqn. (1). The components $v_i$ multiply the base vectors $e_i$ to give three new vectors $v_1e_1$, $v_2e_2$, and $v_3e_3$. The resulting vectors are added together by the head-to-tail rule to give the final vector $v$.

The operation of multiplication of two vectors, say $u$ and $v$, comes in three varieties: The dot product (often called the scalar product) is denoted $u \cdot v$; the cross product (often called the vector product) is denoted $u \times v$; and the tensor product is denoted $u \otimes v$. Each of these products has its own physical significance. In the following sections we review the definitions of these terms, and examine the meaning behind carrying out such operations.

![Figure 3](image.png)

**Figure 3** The components of a vector relative to a basis

![Figure 4](image.png)

**Figure 4** Vector addition and subtraction by the head-to-tail or parallelogram rule
The dot product. The dot product is a scalar value that is related to not only the lengths of the vectors, but also the angle between them. In fact, the dot product can be defined through the formula

$$\mathbf{u} \cdot \mathbf{v} = \| \mathbf{u} \| \| \mathbf{v} \| \cos \theta(\mathbf{u}, \mathbf{v})$$

where $\cos \theta(\mathbf{u}, \mathbf{v})$ is the cosine of the angle $\theta$ between the vectors $\mathbf{u}$ and $\mathbf{v}$, shown in Fig. 5. The definition of the dot product can be expressed directly in terms of the vectors $\mathbf{u}$ and $\mathbf{v}$ by using the law of cosines, which states that

$$\| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 = \| \mathbf{v} - \mathbf{u} \|^2 + 2 \| \mathbf{u} \| \| \mathbf{v} \| \cos \theta(\mathbf{u}, \mathbf{v})$$

Using this result to eliminate $\theta$ from Eqn. (2), we obtain the equivalent definition of the dot product

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2} \left( \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 - \| \mathbf{v} - \mathbf{u} \|^2 \right)$$

We can think of the dot product as measuring the relative orientation between two vectors. The dot product gives us a means of defining orthogonality of two vectors. Two vectors are orthogonal if they have an angle of $\pi/2$ radians between them. According to Eqn. (2), any two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$. If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then they are the legs of a right triangle with the vector $\mathbf{v} - \mathbf{u}$ forming the hypotenuse. In this case, we can see that the Pythagorean theorem makes the right-hand side of Eqn. (3) equal to zero. Thus, $\mathbf{u} \cdot \mathbf{v} = 0$, as before.

Equation (3) suggests a means of computing the length of a vector. The dot product of a vector $\mathbf{v}$ with itself is $\mathbf{v} \cdot \mathbf{v} = \| \mathbf{v} \|^2$. With this observation Eqn. (2) verifies that the cosine of zero (the angle between a vector and itself) is one.

The dot product is commutative, that is, $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$. The dot product also satisfies the distributive law. In particular, for any three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ and scalars $\alpha$, $\beta$, and $\gamma$, we have

$$\alpha \mathbf{u} \cdot (\beta \mathbf{v} + \gamma \mathbf{w}) = \alpha \beta (\mathbf{u} \cdot \mathbf{v}) + \alpha \gamma (\mathbf{u} \cdot \mathbf{w})$$

The dot product can be computed from the components of the vectors as

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{3} u_i e_i \cdot \sum_{j=1}^{3} v_j e_j = \sum_{i=1}^{3} \sum_{j=1}^{3} u_i v_j (e_i \cdot e_j)$$
In the first step we merely rewrote the vectors $u$ and $v$ in component form. In the second step we simply distributed the sums. If the last step puzzles you then you should write out the sums in longhand to demonstrate that the mathematical maneuver was legal. Because the base vectors are orthogonal and of unit length, the products $e_i \cdot e_j$ are all either zero or one. Hence, the component form of the dot product reduces to the expression

$$u \cdot v = \sum_{i=1}^{3} u_i v_i$$  

(5)

The dot product of the base vectors arises so frequently that it is worth introducing a shorthand notation. Let the symbol $\delta_{ij}$ be defined such that

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

(6)

The symbol $\delta_{ij}$ is often referred to as the **Kronecker delta**. Clearly, we can write $e_i \cdot e_j = \delta_{ij}$. When the Kronecker delta appears in a double summation, that part of the summation can be carried out explicitly (even without knowing the values of the other quantities involved in the sum!). This operation has the effect of contraction from a double sum to a single summation, as follows

$$\sum_{i=1}^{3} \sum_{j=1}^{3} u_i v_j \delta_{ij} = \sum_{i=1}^{3} u_i v_i$$

A simple way to see how this contraction comes about is to write out the sum of nine terms and observe that six of them are multiplied by zero because of the definition of the Kronecker delta. The remaining three terms always share a common value of the indices and can, therefore, be written as a single sum, as indicated above.

One of the most important geometric uses of the dot product is the computation of the projection of one vector onto another. Consider a vector $v$ and a unit vector $n$, as shown in Fig. 6. The dot product $v \cdot n$ gives the amount of the vector $v$ that points in the direction $n$. The proof is quite simple. Note that $abc$ is a right triangle. Define a second unit vector $m$ that points in the direction $bc$. By construction $m \cdot n = 0$. Now let the length of side $ab$ be $\gamma$ and the length $\beta$.

![Figure 6](image)  

The dot product gives the amount of $v$ pointing in the direction $n$.
of side bc be $\beta$. The vector $ab$ is then $\gamma n$ and the vector $bc$ is $\beta m$. By the head-to-tail rule we have $v = \gamma n + \beta m$. Taking the dot product of both sides of this expression with $n$ we arrive at the result

$$v \cdot n = (\gamma n + \beta m) \cdot n = \gamma$$

since $n \cdot n = 1$. But $\gamma$ is the length of the side $ab$, proving the original assertion. This observation can be used to show that the dot product of a vector with one of the base vectors has the effect of picking out the component of the vector associated with the base vector used in the dot product. To wit,

$$e_m \cdot v = e_m \cdot \sum_{i=1}^{3} v_i e_i = \sum_{i=1}^{3} v_i \delta_{im} = v_m$$

(7)

We can summarize the geometric significance of the vector components as

$$v_m = e_m \cdot v$$

(8)

That is, $v_m$ is the amount of $v$ pointing in the direction $e_m$.

**The cross product.** The cross product of two vectors $u$ and $v$ results in a vector $u \times v$ that is orthogonal to both $u$ and $v$. The length of $u \times v$ is defined as being equal to the area of a parallelogram, two sides of which are described by the vectors $u$ and $v$. To wit

$$A(u, v) = ||u \times v||$$

(9)

as shown in Fig. 7. The direction of the resulting vector is defined according to the right-hand rule. The cross product is not commutative, but it satisfies the condition of skew symmetry $u \times v = -v \times u$. In other words, reversing the order of the product only changes the direction of the resulting vector. The base vectors satisfy the following identities

$$e_1 \times e_2 = e_3 \quad e_2 \times e_1 = -e_3$$
$$e_2 \times e_3 = e_1 \quad e_3 \times e_2 = -e_1$$
$$e_3 \times e_1 = e_2 \quad e_1 \times e_3 = -e_2$$

(10)

Like the dot product, the cross product is distributive. For any three vectors $u$, $v$, and $w$ and scalars $\alpha$, $\beta$, and $\gamma$, we have

$$u \times v = \alpha (u \times v) + \beta v \times w + \gamma w \times u$$

Figure 7 Area and the cross product of vectors
\[ a\mathbf{u} \times (\beta \mathbf{v} + \gamma \mathbf{w}) = \alpha \beta (\mathbf{u} \times \mathbf{v}) + \alpha \gamma (\mathbf{u} \times \mathbf{w}) \]  

(11)

The component form of the cross product of vectors \( \mathbf{u} \) and \( \mathbf{v} \) is

\[
\mathbf{u} \times \mathbf{v} = \sum_{i=1}^{3} u_i \mathbf{e}_i \times \sum_{j=1}^{3} v_j \mathbf{e}_j = \sum_{i=1}^{3} \sum_{j=1}^{3} u_i v_j (\mathbf{e}_i \times \mathbf{e}_j)
\]

where, again, we have first represented the vectors in component form and then distributed the product. Carrying out the summations, substituting the appropriate incidences of Eqn. (10) for each term of the sum, the component form of the cross product reduces to the expression

\[
\mathbf{u} \times \mathbf{v} = (u_2 v_3 - u_3 v_2) \mathbf{e}_1 + (u_3 v_1 - u_1 v_3) \mathbf{e}_2 + (u_1 v_2 - u_2 v_1) \mathbf{e}_3
\]  

(12)

**The triple scalar product.** The *triple scalar product* of three vectors \( \mathbf{u} \), \( \mathbf{v} \), and \( \mathbf{w} \) is denoted as \((\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}\). Since the dot product results in a scalar and the cross product results in a vector, the order of multiplication is important (and is shown with parentheses). The triple scalar product has an important geometric interpretation. Consider the parallelepiped defined by the three vectors \( \mathbf{u} \), \( \mathbf{v} \), and \( \mathbf{w} \) shown in Fig. 8. The cross product of \( \mathbf{u} \) and \( \mathbf{v} \) results in a vector that is normal to both \( \mathbf{u} \) and \( \mathbf{v} \). Let us normalize this vector by its length to define the unit vector \( \mathbf{n} \equiv \mathbf{u} \times \mathbf{v} / \| \mathbf{u} \times \mathbf{v} \| \). The height of the parallelepiped perpendicular to its base is the length of the component of \( \mathbf{w} \) that lies along the unit vector \( \mathbf{n} \). This height is simply \( h = \mathbf{w} \cdot \mathbf{n} \). Thus, the volume of the parallelepiped is the base area times the height

\[
V(\mathbf{u}, \mathbf{v}, \mathbf{w}) = h A(\mathbf{u}, \mathbf{v}) = \left( \mathbf{w} \cdot \frac{\mathbf{u} \times \mathbf{v}}{\| \mathbf{u} \times \mathbf{v} \|} \right) \| \mathbf{u} \times \mathbf{v} \|
\]

Upon simplification, we get the following formula for the volume of the parallelepiped as the triple scalar product of the three vectors \( \mathbf{u}, \mathbf{v}, \) and \( \mathbf{w} \)

\[
V(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}
\]

(13)

The triple scalar product can be computed in terms of components. Taking the dot product of \( \mathbf{w} \) with \( \mathbf{u} \times \mathbf{v} \), as already given in Eqn. (12), we find

![Figure 8](image-url)
\[ (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = w_1(u_2v_3 - u_3v_2) + w_2(u_3v_1 - u_1v_3) + w_3(u_1v_2 - u_2v_1) \]

\[ = (u_1v_2w_3 + u_2v_3w_1 + u_3v_1w_2) - (u_3v_2w_1 + u_2v_1w_3 + u_1v_3w_2) \]

where the second form shows quite clearly that the indices are distinct for each term and that the indices on the positive terms are in \textit{cyclic} order while the indices on the negative terms are in \textit{acyclic} order. Cyclic and acyclic order can be easily visualized, as shown in Fig. 9. If the numbers 1, 2, and 3 appear on a circle in clockwise order, then a cyclic permutation is the order in which you encounter these numbers when you move clockwise from any starting point, and an acyclic permutation is the order in which you encounter them when you move anticlockwise. The indices are in cyclic order when they take the values (1, 2, 3), (2, 3, 1), or (3, 1, 2). The indices are in acyclic order when they take the values (3, 2, 1), (1, 3, 2), or (2, 1, 3).

\[ \begin{array}{c}
1 \\
3 \\
2 \\
\end{array} \quad \begin{array}{c}
1 \\
3 \\
2 \\
\end{array} \\
\text{Cyclic} \quad \text{Acyclic}
\]

\[ \text{Figure 9} \quad \text{Cyclic and acyclic permutations of the numbers 1, 2, and 3} \]

The triple scalar product of base vectors represents a fundamental geometric quantity. It will be used in Chapter 2 to describe the volume of a solid body and the changes in that volume. Let us introduce a shorthand notation that is related to the triple scalar product. Let the (permutation) symbol \( \epsilon_{ijk} \) be

\[
\epsilon_{ijk} \equiv \begin{cases} 
1 & \text{if } (i, j, k) \text{ are in cyclic order} \\
0 & \text{if any of } (i, j, k) \text{ are equal} \\
-1 & \text{if } (i, j, k) \text{ are in acyclic order} 
\end{cases} \tag{14}
\]

The scalars \( \epsilon_{ijk} \) are sometimes referred to as the components of the \textit{permutation tensor}. There are 27 possible permutations of three indices that can each take on three values. Of these 27, only three have (distinct) cyclic values and only three have (distinct) acyclic values. All other permutations of the indices involve equality of at least two of the indices. The 27 possible values of the permutation symbol can be summarized with the triple scalar products of the base vectors. To wit,

\[
(\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_k = \epsilon_{ijk} \tag{15}
\]

With the permutation symbol, the cross product and the triple scalar product can be expressed neatly in component form as
Chapter 1 Vectors and Tensors

\[ \mathbf{u} \times \mathbf{v} = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} u_i v_j \epsilon_{ijk} e_k \]  
\[ (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} u_i v_j w_k \epsilon_{ijk} \]  

(16)

You should verify that these formulas involving \( \epsilon_{ijk} \) give the same results as found previously.

Tensors

The cross product is an example of a vector operation that has as its outcome a new vector. It is a very special operator in the sense that it produces a vector orthogonal to the plane containing the two original vectors. There is a much broader class of operations that produce vectors as the result. The second-order tensor is the mathematical object that provides the appropriate generalization. (If the context is not ambiguous, we will often refer to a second-order tensor simply as a tensor.)

**Definition.** A tensor is an object that operates on a vector to produce another vector.

(17)

Schematically, this operation is shown in Fig. 10, wherein a tensor \( T \) operates on the vector \( \mathbf{v} \) to produce the new vector \( T\mathbf{v} \). Unlike a vector, there is no easy graphical representation of the tensor \( T \) itself. In abstract we shall understand a tensor by observing what it does to a vector. The example shown in Fig. 10 is illustrative of all tensor actions. The vector \( \mathbf{v} \) is stretched and rotated to give the new vector \( T\mathbf{v} \). In essence, tensors stretch and rotate vectors.

A tensor is a linear operator that satisfies

\[ T(\alpha \mathbf{u} + \beta \mathbf{v} + \gamma \mathbf{w}) = \alpha T\mathbf{u} + \beta T\mathbf{v} + \gamma T\mathbf{w} \]  

(18)

for any three scalars \( \alpha, \beta, \gamma \), and any three vectors \( \mathbf{u}, \mathbf{v}, \mathbf{w} \). Because any vector in three-dimensional space can be expressed as a linear combination of three vectors that span the space, it is sufficient to consider the action of the tensor on three independent vectors. The action of the tensor \( T \) on the base vectors, for example, completely characterizes the action of the tensor on any other vector. Thus, it is evident that a tensor can be completely characterized by nine

Figure 10 A tensor operates on a vector to produce another vector
scalar quantities: the three components of the vector $T e_1$, the three components of the vector $T e_2$, and the three components of the vector $T e_3$. We shall refer to these nine scalar quantities as the components of the tensor. Like a vector, which can be expressed as the sum of scalar components times base vectors, we shall represent a tensor as the sum of scalar components times base tensors. We introduce the tensor product of vectors as the building block to define a natural basis for a second-order tensor.

The tensor product of vectors. The tensor product of two vectors $u$ and $v$ is a special second-order tensor which we shall denote $[u \otimes v]$. The action of this tensor is embodied in how it operates on a vector $w$, which is

$$[u \otimes v] w = (v \cdot w) u$$

(19)

In other words, when the tensor $u \otimes v$ operates on $w$ the result is a vector that points in the direction $u$ and has the length equal to $(v \cdot w) \|u\|$, the original length of $u$ multiplied by the scalar product of $v$ and $w$. The tensor product of vectors appears to be a rather curious object, and it certainly takes some getting used to. It will, however, prove to be highly useful in developing a coordinate representation of a general tensor $T$.

The tensor products of the base vectors $e_i \otimes e_j$ comprise a set of second-order tensors. Since there are three base vectors, there are nine distinct tensor product combinations among them. These nine tensors provide a suitable basis for expressing the components of a tensor, much like the base vectors themselves provided a basis for expressing the components of a vector. Like the base vectors, we presume to understand these base tensors better than any other tensors in the space. We can confirm that by noting that their action is given simply by Eqn. (19). In fact, we can observe from Eqn. (19) that

$$[e_i \otimes e_j] e_k = (e_j \cdot e_k) e_i = \delta_{ik} e_i$$

(20)

We will use this knowledge of the tensor product of base vectors to help us with the manipulation of tensor components.

The second-order tensor $T$ can be expressed in terms of its components $T_{ij}$ relative to the base tensors $e_i \otimes e_j$ as

$$T = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} [e_i \otimes e_j]$$

(21)

It will soon be evident why we elect to represent the nine scalar components with a double indexed quantity. Like vector components, the components $T_{ij}$ are scalar values that depend upon the basis chosen for the representation. The tensor part of $T$ comes from the base tensors $e_i \otimes e_j$. The tensor, then, is a sum of scalars times base tensors. Like a vector, the tensor $T$ itself does not depend upon the coordinate system; only the components do.
A tensor is completely characterized by its action on the three base vectors. Let us compute the action of $T$ on the base vector $e_n$

$$Te_n = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij}[e_i \otimes e_j]e_n = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij}\delta_{jn}e_i = \sum_{i=1}^{3} T_{in}e_i \quad (22)$$

The first step simply introduces the coordinate form of $T$. The second step carries out the tensor product of vectors as in Eqn. (20). The final step recognizes that the sum of nine terms reduces to a sum of three terms because six of the nine terms are equal to zero.

We can get some insight into the physical significance of the components by taking the dot product of $e_m$ and $Te_n$. Recall from Eqn. (8) that dotting a vector with $e_m$ simply extracts the $m$th component of the vector. Starting from the result of Eqn. (22) we compute

$$e_m \cdot Te_n = e_m \cdot \sum_{i=1}^{3} T_{in}e_i = \sum_{i=1}^{3} T_{in}\delta_{im} = T_{mn} \quad (23)$$

Thus, we can see that $T_{mn}$ is the $m$th component of the vector $Te_n$. We can summarize the physical significance of the tensor components as follows

$$T_{mn} = e_m \cdot Te_n \quad (24)$$

**The identity tensor.** The identity tensor is the tensor that has the property of leaving a vector unchanged. We shall denote the identity tensor as $I$, and endow it with the property that $Iv = v$, for all vectors $v$. The identity tensor can be expressed in terms of orthonormal (i.e., orthogonal and unit) base vectors

$$I = \sum_{i=1}^{3} e_i \otimes e_i \quad (25)$$

Of course, this definition holds for any orthonormal basis. To prove that Eqn. (25), we need only consider the action of $I$ on a base vector $e_j$. To wit

$$Ie_j = \sum_{i=1}^{3} [e_i \otimes e_i]e_j = \sum_{i=1}^{3} (e_i \cdot e_j) e_i = \sum_{i=1}^{3} \delta_{ij}e_i = e_j$$

Since the base vectors span three-dimensional space, it is apparent that $Iv = v$ for any vector. Observe that Eqn. (25) can be expressed in terms of the Kronecker delta as

$$I = \sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij}[e_i \otimes e_j]$$
Hence, $\delta_{ij}$ can be interpreted as the $ij$th component of the identity tensor.

**The tensor inverse.** Let us assume that we have a tensor $T$ and that it acts on a vector $v$ to produce another vector $Tv$. A tensor stretches and rotates a vector. It seems reasonable to imagine a tensor that undoes the action of another tensor. Such a tensor is called the *inverse* of the tensor $T$, and we denote it as $T^{-1}$. Thus, $T^{-1}$ is the tensor that exactly undoes what the tensor $T$ does. To be more specific, the tensor $T^{-1}$ can be applied to the vector $Tv$ to give back $v$. Conversely, if the tensor $T^{-1}$ is applied to the vector $v$ to give the vector $T^{-1}v$, then the tensor $T$ can be applied to $T^{-1}v$ to give back the vector $v$. These operations define the inverse of a tensor and are summarized as follows

$$T^{-1}(Tv) = v, \quad T(T^{-1}v) = v$$  \hspace{1cm} (26)

The above relations hold for any vector $v$. As we will soon see, the composition of tensors (a tensor operating on a tensor) can be viewed as a tensor itself. Thus, we can say that $T^{-1}T = I$ and $TT^{-1} = I$.

**Example 1.** As a simple example of a tensor and its operation on vectors, consider the *projection* tensor $P$ that generates the image of a vector $v$ projected onto the plane with normal $n$, as shown in Fig. 11.

![Figure 11](image)

**Figure 11** The action of the projection tensor

The explicit expression for the tensor is given by

$$P = I - n \otimes n$$  \hspace{1cm} (27)

where $I$ is the identity tensor. The action of $P$ on $v$ gives the result

$$Pv = \left[I - n \otimes n\right]v = Iv - [n \otimes n]v = v - (n \cdot v)n$$

To see that the vector $Pv$ lies in the plane we need only to show that its dot product with the normal vector $n$ is zero. Accordingly, we can make the computation $Pv \cdot n = (v \cdot n) - (v \cdot n)(n \cdot n) = 0$, since $n$ is a unit vector.

It is interesting to note that we can derive the tensor $P$ from geometric considerations. From Fig. 11 we can see that, by vector addition, $Pv + \beta n = v$ for some, as yet unknown, value of the scalar $\beta$. To determine $\beta$ we simply take the dot product of the previous vector equation with the vector $n$, noting that $n$ has unit length and is perpendicular to $Pv$. Hence, $\beta = v \cdot n$. Now, we substitute back to get
Chapter 1  Vectors and Tensors

\[ \mathbf{Pv} = \mathbf{v} - \beta \mathbf{n} = \mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} = [\mathbf{I} - \mathbf{n} \otimes \mathbf{n}] \mathbf{v} \]  

(28)

thereby determining the tensor \( \mathbf{P} \).

**Component expression for operation of a tensor on a vector.** Equipped with the component representation of a tensor we can now take another look at how a tensor \( \mathbf{T} \) operates on a vector \( \mathbf{v} \). In particular, let us examine the components of the resulting vector \( \mathbf{Tv} \).

\[
\mathbf{Tv} = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} [\mathbf{e}_i \otimes \mathbf{e}_j] \sum_{k=1}^{3} \mathbf{v}_k \mathbf{e}_k = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} T_{ij} \mathbf{v}_k [\mathbf{e}_i \otimes \mathbf{e}_j] \mathbf{e}_k
\]

(29)

Carrying out the summations in Eqn. (29), noting the properties expressed in Eqn. (20), we finally obtain the result

\[
\mathbf{Tv} = \sum_{i=1}^{3} \sum_{j=1}^{3} T_{ij} \mathbf{v}_j \mathbf{e}_i
\]

(30)

From this expression, we can see that the result is a vector (anything expressed in a vector basis is a vector). Furthermore, we can observe from Eqn. (30) that the \( i \)th component of the vector \( \mathbf{Tv} \) is given by

\[
(Tv)_i = \sum_{j=1}^{3} T_{ij} \mathbf{v}_j
\]

(31)

That is, we compute the \( i \)th component of the resulting vector from the components of the tensor and the components of the original vector. The similarity between the operation of a tensor and that of a matrix in linear algebra should be apparent.

**The summation convention.** General relativity is a theory based on tensors. While Einstein was working on this theory, he apparently got rather tired of writing the summation symbol with its range of summation decorating the bottom and top of the Greek letter sigma. What he observed was that, most of the time, the range of the summation was equal to the dimension of space (three dimensions for us, four for him) and that when the summation involved a product of two terms, the summation was over a repeated index. For example, in Eqn. (31) the index \( j \) is the index of summation, and it appears exactly twice in the summand \( T_{ij} \mathbf{v}_j \). Einstein decided that, with a little care, summations could be expressed without laboriously writing the summation symbol. The summation symbol would be understood to apply to repeated indices.

The summation convention, then, means that any repeated index, also called a **dummy index**, is understood to be summed over the range 1 to 3. With the summation convention, then, Eqn. (30) can be written as
Tv = T^vje, with the summation on the indices i and j implied because both are repeated. All we have done is to eliminate the summation symbol, a pretty significant economy of notation. The triple scalar product of vectors can now be written

\[(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = u_i v_j w_k \varepsilon_{ijk}\]

Indices that are not repeated in a product are called free indices. These indices are not summed and must appear on both sides of the equation. For example, the index i in the equation

\[(\mathbf{Tv})_i = T^i_j v_j\]

is a free index. The presence of free indices really indicate multiple equations. The index equation must hold for all values of the free index. The equation above is really three equations,

\[(\mathbf{Tv})_1 = T^1_j v_j, \quad (\mathbf{Tv})_2 = T^2_j v_j, \quad (\mathbf{Tv})_3 = T^3_j v_j\]

That is, the free index i takes on values 1, 2, and 3, successively.

The letter used for a dummy index can be changed at will without changing the value of the expression. For example,

\[(\mathbf{Tv})_i = T^i_j v_j = T^i_k v_k\]

A free index can be renamed if it is renamed on both sides of the equation. The previous equation is identical to

\[(\mathbf{Tv})_m = T^m_j v_j = T^m_k v_k\]

The beauty of this shorthand notation should be apparent. But, like any notational device it should be used with great attention to detail. The mere slip of an index can ruin a derivation or computation.

Perhaps the greatest pitfall of the novice index manipulator is to use an index too many times. An expression with an index appearing more than twice is ambiguous and, therefore, meaningless. For example, the term T^i_j v_j has no meaning because the summation is ambiguous. The summation convention applies only to terms involved in the same product; to indices of the same tensor, as in the case T^i_j = T^11_{ij} + T^22_{ij} + T^33_{ij}; and to indices in a quotient, as in the expression for divergence, i.e., \(\frac{\partial V_1}{\partial x_1} + \frac{\partial V_2}{\partial x_2} + \frac{\partial V_3}{\partial x_3}\). Terms separated by a + operation are not subject to the summation convention, and in such a case an index can be reused, as in the expression \(T^i_j v_j + S^i_j w_j\). Whenever the Kronecker delta appears in a summation, it has the net effect of contracting indices. For example

\[T^i_j \delta_{jk} = T^i_k\]
Observe how the summed index \( j \) on the tensor component \( T_{ij} \) is simply replaced by the free index \( k \) on \( \delta_{jk} \) in the process of contraction of indices.

In this book the summation convention will be in force unless specifically indicated otherwise.

**Generating tensors from other tensors.** We can define sums and products of tensors using only the geometric and operational notions of vector addition and multiplication. For example, we know how to add two vectors so that the operation \( T_v + S_v \) makes sense (by the head-to-tail rule). The question is: Does the operation \( T + S \) make sense? In other words, can you add two tensors together? It makes sense if we define it to make sense. So we will.

Let us define the sum of two tensors \( T \) and \( S \) through the following operation

\[
[T + S] v = T v + S v
\]

In other words, the tensor \([T + S]\) operating on a vector \( v \) is equivalent to the sum of the vectors created by \( T \) and \( S \) individually operating on the vector \( v \).

An expression for the components of the tensor \([T + S]\) can then be constructed simply using the component expressions for Eqn. (32). Let us use Eqn. (30), which gives the formula for computing the components of a tensor operating on a vector, as the starting point (no need to reinvent the wheel). We can write each term of Eqn. (32) in component form and then gather terms on the right side of the equation to yield

\[
[T + S]_{ij} v^j e_i = T_{ij} v^j e_i + S_{ij} v^j e_i = (T_{ij} + S_{ij}) v^j e_i
\]

From simple identification of terms on both sides of the equation, we get

\[
[T + S]_{ij} = T_{ij} + S_{ij}
\]

In other words, the \( ij \)th component of the sum of two tensors is the sum of the \( ij \)th components of the two original tensors.

We can follow the same approach to define multiplication of a tensor by a scalar, as in \( \alpha T \). The scaled tensor \( \alpha T \) is defined through the operation

\[
[\alpha T] v = \alpha (T v)
\]

Again, the component expression can be deduced by applying Eqn. (30) to get

\[
[\alpha T]_{ij} v^j e_i = \alpha (T_{ij} v^j e_i) = (\alpha T_{ij}) v^j e_i
\]

Thus, the components of the scaled tensor are \([\alpha T]_{ij} = \alpha T_{ij}\). That is, each component of the original tensor is scaled by \( \alpha \).
The definition of the **transpose** of a tensor can be constructed as follows. The dot product \( u \cdot Tv \) is a scalar. One might wonder if there is a tensor for which we could reverse the order of operation on \( u \) and \( v \) and get exactly the same scalar value. There is and the tensor is called the transpose of \( T \). We shall use the symbol \( T' \) to denote the transpose. The transpose of \( T \) is defined through the identity

\[
v \cdot T'u = u \cdot Tv
\]  

(34)

The components of the transpose \( T' \) can be shown to be \([T']_i^j = [T]_j^i\) (see Problem 10). That is, the first and second index (row and column in matrix notation) of the tensor components are simply swapped. A tensor is called symmetric if the operation of the tensor and its transpose give identical results, i.e., \( u \cdot Tv = v \cdot Tu \). The components of a symmetric tensor satisfy \( T_{ij} = T_{ji} \).

We can define a new tensor through the composition of two tensors \([ST]\). Let the tensor \( S \) operate on the vector \( Tv \). We can define the tensor \([ST]\) as

\[
[ST]v = S(Tv)
\]  

(35)

The components of the tensor \( ST \) can be computed as follows

\[
[ST]_{i}^{j}e_{i} = S_{\alpha}[e_{j} \otimes e_{\alpha}](T_{mj}e_{m})
\]

\[
= (S_{\alpha}T_{mj}\delta_{km})v_{j}e_{i}
\]

Contracting the index \( m \) in the above expression leads to the formula for the components of the composite tensor

\[
[ST]_{ij} = S_{\alpha}T_{ij}
\]  

(36)

Notice how close is the resemblance between this formula and the formula for the product of two square matrices.

An alternative composition of two second-order tensors can also be defined using the dot product of vectors. Consider two tensors \( S \) and \( T \). Let the two tensors operate on the vectors \( u \) and \( v \) to give two new vectors \( Su \) and \( Tv \). Now we can take the dot product of the new vectors. According to Eqn. (34), this product is equal to

\[
Su \cdot Tv = u \cdot S'(Tv) = u \cdot [S'T]v
\]

We can view the tensor \( S'T \) as a second-order tensor in its own right, operating on the vector \( v \) and then dotted with \( u \). The tensor \( S'T \) has components

\[
[S'T]_{ij} = S_{\alpha}T_{ij}
\]  

(37)

Notice the subtle difference between Eqns. (36) and (37). The tensor \( T' \) is always symmetric, even if \( T \) is not (see Problem 11).
It should be clear that we could go on defining new tensor objects ad infinitum. Any such definition will emanate from the same basic considerations, and the computation of the components of the resulting tensors follows exactly along the lines given above. We shall have the opportunity to make such definitions throughout this book, and thus defer further discussion until needed.

**Tensors, tensor components, and matrices.** A tensor is not a matrix. However, if the foregoing discussion of tensors has left you thinking of matrices, you are not far off the mark. The way we have chosen to denote the components of a second-order tensor (with two indices, that is) makes the temptation to think of tensors as matrices quite compelling. We can list the components of a tensor in a matrix; all of the formulas for tensor-index manipulation are then exactly the same as standard matrix algebra. To some extent, matrix algebra can be an aid to understanding formulas like Eqn. (36). On the other hand, a second-order tensor is no more a three by three matrix than a vector is a three by one matrix.

Matrices are for keeping books, for organizing computations. A tensor or a vector exists independent of a particular manifestation of its components; a matrix is a particular manifestation of its components. So take the analogy between tensors and matrices for what it is worth, but try not to confuse a tensor with its components. To do so is rather like being unable to feel cold because you don’t know the value of the temperature in degrees Celsius. The fundamental property of “cold” exists independent of what scale you choose to measure temperature.

That said, let us back off from this purist view a little and introduce a notational shorthand that will be useful in stating and solving problems in tensor analysis. When we solve a particular problem, we will select a coordinate system having a particular set of base vectors. The components of any tensor will be expressed relative to those base vectors. For expedience, we will often collect those components in a matrix as

\[
T \sim \begin{bmatrix}
T_{11} & T_{12} & T_{13} \\
T_{21} & T_{22} & T_{23} \\
T_{31} & T_{32} & T_{33}
\end{bmatrix}
\]

where the notation \( T \sim [ \ ] \) should be read as “the components of the tensor \( T \), relative to the understood basis, are stored in the matrix \([ \ ]\) with the convention that the first index \( i \) on the tensor component \( T_{ij} \) is the row index of the matrix and the second index \( j \) on the tensor component is the column index of the matrix.” We avoid the temptation to use the notation \( T = [ \ ] \) because we do not want to give the impression that we are setting a tensor equal to a matrix of its components. If there is any question as to what the basis is, then this abbreviated notation does not make sense, and should not be used. The reason this
notation is useful is that tensor multiplication is the same as matrix multiplication if the components are stored in the manner shown.

**Change of basis.** Consider two different coordinate systems, the first with unit base vectors \( \{ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \} \) and the second with unit base vectors \( \{ \mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3 \} \). Any vector \( \mathbf{v} \) can be expressed in terms of its components along the base vectors of a coordinate system, as shown in Fig. 12. Clearly, the components of a vector depend upon the coordinate system even though the vector itself does not. It seems reasonable that the components of the vector with respect to one basis should be related somehow to the components of the vector with respect to the other basis. In this section we shall derive that relationship.

A vector can be expressed equivalently in the two bases as

\[
\mathbf{v} = v_j \mathbf{e}_j = \hat{v}_j \mathbf{g}_j
\]  

We can derive the relationship between the two sets of components by taking the dot product of the vector \( \mathbf{v} \) with one of the base vectors, say \( \mathbf{g}_i \). From Eqn. (38) we obtain

\[
\mathbf{g}_i \cdot \mathbf{v} = \hat{v}_i = v_j (\mathbf{g}_i \cdot \mathbf{e}_j)
\]

since \( \hat{v}_j (\mathbf{g}_i \cdot \mathbf{e}_j) = \hat{v}_j \delta_{ij} = \hat{v}_i \). Let us define the nine scalar values

\[
Q_{ij} \equiv \mathbf{g}_i \cdot \mathbf{e}_j
\]

that arise from the dot products of the base vectors. The nine values record the cosines of the angles between the nine pairings of the base vectors. Note that the first index of \( Q \) is associated with the \( \mathbf{g} \) base vector and the second index of \( Q \) is associated with the \( \mathbf{e} \) base vector. Be careful. The dot product is commutative so \( Q_{ij} = Q_{ji} \) (the first index of \( Q \) is still associated with \( \mathbf{g} \) and the second index is still associated with \( \mathbf{e} \)).

The formula giving one set of vector components in terms of the other is then

\[
\hat{v}_i = Q_{ij} v_j
\]
Chapter 1 Vectors and Tensors

We can find the reverse relationship by dotting Eqn. (38) with \( e_i \) instead of \( g_i \). Carrying out a similar calculation we find that

\[
\nu_i = Q_{ij} \nu_j
\]

(41)

The components of a second-order tensor \( T \) transform in a manner similar to vectors. A tensor can be expressed in terms of components relative to two different bases in the following manner

\[
T = \hat{T}_{ij} [g_i \otimes g_j] = T_{ij} [e_i \otimes e_j]
\]

where \( \hat{T}_{ij} \) is the \( ij \)th component of \( T \) with respect to the base tensor \([g_i \otimes g_j]\) and \( T_{ij} \) is the \( ij \)th component of \( T \) with respect to the base tensor \([e_i \otimes e_j]\). The relationship between the components in the two coordinate systems can be found by computing the product \( g^i \cdot T g^j \), as follows

\[
g^i \cdot T g^j = T_{ij} (g^i \cdot e_j) (g^j \cdot e_i)
\]

Computing instead \( e_m \cdot T e_n \), we can find the inverse relationship. Once again noting that \( Q_{ij} \equiv g_i \cdot e_j \), we can write the formulas for the transformation of second-order tensor components as

\[
\hat{T}_{mn} = Q_{mj} T_{nj} \quad T_{mn} = Q_{im} Q_{jn} \hat{T}_{ij}
\]

(42)

The main difference between transforming the components of a tensor and those of a vector is that it took two \( Q \) terms to accomplish the task for a tensor, one for each index, but only one \( Q \) term for a vector. It should be evident that higher-order tensors, i.e., those with more indices, will transform analogously with the appropriate number of \( Q \) terms present.

As you might expect, the components of the coordinate transformation \( Q_{ij} \equiv g_i \cdot e_j \) have some interesting properties. These components make up what is called an orthogonal transformation. The orthogonal transformation components have the following property

\[
Q_{ki} Q_{lj} = \delta_{ij} \quad Q_{ik} Q_{jk} = \delta_{ij}
\]

(43)

The proof of each equation relies on the expression for the identity tensor:

\[
(g_i \cdot e_i)(g_k \cdot e_j) = e_i \cdot [g_k \otimes g_i] e_j = e_i \cdot e_j = \delta_{ij}
\]

\[
(g_i \cdot e_k)(g_j \cdot e_k) = g_i \cdot [e_k \otimes e_k] g_j = g_i \cdot g_j = \delta_{ij}
\]

Problem 13 asks you to explore further the relationship between the two bases, and clarifies the notion of the \( Q_{ij} \) being components of a tensor \( Q \).

Example 2. There is a relationship between the permutation symbol and the Kronecker delta that is often referred to as the \( \epsilon - \delta \) identity. The identity is
Let us prove this identity.

First note that the cross product is equivalent to operation by a skew-symmetric tensor \([u \times]\) defined to have components as follows

\[
[u \times] = \begin{bmatrix}
0 & -w_3 & w_2 \\
-w_3 & 0 & -w_1 \\
w_2 & w_1 & 0
\end{bmatrix}
\]

One can easily verify that \([u \times]\)[v] = u \times v. By matrix multiplication one can also verify that \([u \times]^[v \times] = (u \cdot v)I - v \otimes u\). Now,

\[
\epsilon_{ijk}\epsilon_{imn} = \left( (e_i \times e_j) \cdot e_k \right) \left( (e_i \times e_m) \cdot e_n \right)
\]

\[
= \left( (e_k \times e_j) \cdot e_i \right) \left( e_i \cdot (e_m \times e_n) \right)
\]

\[
= (e_k \times e_j) \cdot \left[ e_i \otimes e_i \right] (e_m \times e_n)
\]

\[
= \left[ e_k \times \left( e_j \cdot e_m \right) \right] e_n
\]

\[
= e_j \cdot \left[ e_k \times \left( e_m \cdot e_n \right) \right] e_k
\]

\[
= e_j \cdot \left( (e_k \cdot e_m)I - e_n \otimes e_k \right) e_n
\]

\[
= (e_j \cdot e_m)(e_k \cdot e_n) - (e_j \cdot e_n)(e_k \cdot e_m)
\]

\[
= \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}
\]

There are other, possibly simpler proofs of the \(\epsilon - \delta\) identity. For example, one can recognize that the identity is simply 81 equations. You can verify them one by one. This example has the additional merit of illustrating various vector and tensor manipulation techniques.

**Tensor invariants.** In subsequent chapters we will have occasions to wonder whether there are properties of the tensor components that do not depend upon the choice of basis. These properties will be called tensor invariants. The identities of Eqn. (43) will be useful in proving the invariance of these properties. The argument will go something like this: Let \(f(T_{ij})\) be a function of the components of the tensor \(T\). Under a change of basis, we can write this function in the form \(f(Q_{ik}Q_{jl}T_{kl})\). If the function has the property that

\[
f(Q_{ik}Q_{jl}T_{kl}) = f(T_{ij})
\]

then the function \(f\) is a tensor invariant. Since it does not depend upon the coordinate system, we can say that it is an intrinsic function of the tensor \(T\), and write \(f(T)\). Three fundamental tensor invariants are given by

\[
\begin{align*}
\epsilon_{ij}(T) & \equiv T_{ii} \\
f_{2}(T) & \equiv T_{ij}T_{ji} \\
f_{3}(T) & \equiv T_{ij}T_{jk}T_{ki}
\end{align*}
\]
The proof that $f_1(T)$ is invariant is straightforward

$$f_1(T) = \hat{T}_{ii} = \delta_{ii} T_{ii} = T_{kk}$$

by the formula for change of basis, contracted to give $\hat{T}_{ii}$, and Eqn. (43). The invariance of the other two functions can be proved in a similar manner (see Problem 18). Any function of tensor invariants is itself a tensor invariant. We shall sometimes refer to the invariant functions $f_1(T)$, $f_2(T)$, and $f_3(T)$ as the primary invariants to distinguish them from other invariant functional forms.

The trace of a tensor is simply the sum of its diagonal components. We use the operator "tr" to designate the trace. Thus, $\text{tr}(T) = T_{ii}$ is the first invariant of the tensor $T$. The second and third invariants can also be expressed in terms of the trace operator. Let us introduce the notation of a tensor raised to a power as $T^2 = TT$ and $T^3 = TTT$, where the components are given by the formula for products of tensors, Eqn. (36), as

$$[T^2]_{ij} = T_{im} T_{mj} \quad [T^3]_{ij} = T_{im} T_{mn} T_{nj}$$

(45)

It should be evident that a tensor can be raised to any (integer) power. Taking the trace of $T^2$ and $T^3$ gives $\text{tr}(T^2) = [T^2]_{ii}$ and $\text{tr}(T^3) = [T^3]_{ii}$. Using these expressions in Eqn. (45) we find that the three invariants can be equivalently cast in terms of traces of powers of the tensor $T$ as

$$f_1(T) = \text{tr}(T), \quad f_2(T) = \text{tr}(T^2), \quad f_3(T) = \text{tr}(T^3)$$

(46)

By extension, one can establish that $f_n(T) = \text{tr}(T^n)$ is an invariant of the tensor $T$ for any value of $n$ (see Problem 18). One can prove that the invariants for $n \geq 4$ can all be computed from the first three invariants (see Problem 19).

**Eigenvalues and eigenvectors of symmetric tensors.** A tensor has properties independent of any basis used to characterize its components. As we have just seen, the components themselves have mysterious properties called invariants that are independent of the basis that defines them. It seems reasonable to expect that we might be able to find a representation of a tensor that is canonical. Indeed, this canonical form is the spectral representation of the tensor that can be built from its eigenvalues and eigenvectors. In this section we shall build the mathematics behind the spectral representation of tensors.

Recall that the action of a tensor is to stretch and rotate a vector. Let us consider a symmetric tensor $T$ acting on a unit vector $n$. If the action of the tensor is simply to stretch the vector but not to rotate it then we can express it as

$$Tn = \mu n$$

(47)

where $\mu$ is the amount of the stretch. This equation, by itself, begs the question of existence of such a vector $n$. Is there any vector that has the special property
that action by \( T \) is identical to multiplication by a scalar? Is it possible that more than one vector has this property?

Equation (47) is called an eigenvalue problem. Eigenvalue problems show up all over the place in mathematical physics and engineering. The tensor in three dimensional space is a great context in which to explore the eigenvalue problem because the computations are quite manageable (as opposed to, say, solving the vibration eigenvalue problem of structural dynamics on a structure with a million degrees of freedom).

A vector \( n \) that satisfies the eigenvalue problem is a special vector (an eigenvector) that has the property that operation by the second-order tensor \( T \) is the same as operation by the scalar \( \mu \) (the eigenvalue). Equation (47) can be written as \([T - \mu I]n = 0\), which is a linear homogeneous system of equations. (Note that 0 is the zero vector). In order for this system to have a nontrivial solution (i.e., \( n \neq 0 \)), the determinant of the coefficient matrix must be equal to zero. That is,

\[
\det[T - \mu I] = \det \begin{bmatrix}
T_{11} - \mu & T_{12} & T_{13} \\
T_{21} & T_{22} - \mu & T_{23} \\
T_{31} & T_{32} & T_{33} - \mu \\
\end{bmatrix} = 0 \quad (48)
\]

If we carry out the computation of the determinant, we get the characteristic equation (a cubic equation in the case of a three by three matrix) for the eigenvalues \( \mu \). The characteristic equation can be written in the form

\[
- \mu^3 + I_T \mu^2 - II_T \mu + III_T = 0 \quad (49)
\]

where the coefficients of the characteristic polynomial

\[
I_T = \text{tr}(T), \quad II_T = \frac{1}{2}[I_T - \text{tr}(T^2)], \quad III_T = \det(T) \quad (50)
\]

are invariants of the tensor \( T \). We shall refer to \( I_T, II_T \), and \( III_T \) as the principal invariants to distinguish these functions from the primary invariants. The determinant of a tensor can be expressed in terms of the primary invariants \( f_1(T) \), \( f_2(T) \), and \( f_3(T) \) (see Problem 23), so all three of the principal invariants are functions of the primary invariants (and vice versa). The principal invariants can be expressed in component form as

\[\dagger\] The definition of the eigenvalue problem does not require that \( n \) be a unit vector. In fact, it should be obvious that if \( n \) satisfies Eqn. (47) then so does any scalar multiple of \( n \). Setting the length of the eigenvector is usually considered arbitrary with many choices available. However, in many applications there is an auxiliary condition that determines the length of the vector. For the two most important cases that we will consider in solid mechanics (principal values of stress and strain tensors) the vector \( n \) must be unit length. Assuming unit length from the outset removes some ambiguity without loss of generality.
Because the coefficients of the characteristic equation are invariants of the tensor \( T \) it follows that the roots \( \mu \) do not depend upon the basis chosen to describe the components and hence are intrinsic properties of \( T \).

**Finding the roots of the characteristic equation.** The cubic equation has three roots (not necessarily distinct) that correspond to three (not necessarily unique) directions. If the cubic equation cannot be factored, then the roots can be found iteratively. For example, we can use Newton's method to solve the nonlinear equation \( g(x) = 0 \). Given a starting value \( x_0 \), we can compute successive estimates of a root of \( g(x) = 0 \) (see Chapter 12) as

\[
x_{i+1} = x_i - \frac{g(x_i)}{g'(x_i)}
\]

where \( g'(x_i) \) is the derivative of \( g(x) \) evaluated at the current iterate \( x_i \). The starting value determines the root to which the iteration converges if there are multiple roots. In the present context, let \( x_i \) be the estimate of the eigenvalue \( \mu \) at the \( i \)th iteration. The next estimate can be computed from Newton's formula as

\[
x_{i+1} = \frac{2x_i^3 - I_T x_i^2 + III_T}{3x_i^2 - 2I_T x_i + II_T}
\]

The iteration continues until \( |x_n - x_{n-1}| \) is less than some acceptable tolerance. Then the eigenvalue is \( \mu = x_n \).

We can always take, as a starting value, \( x_0 = 0 \). However, Gershgorin's theorem might be of some help in estimating a good starting point for the Newton iteration. Gershgorin's theorem simply states that the diagonal element \( T_{ii} \) of the tensor \( T \) might be a good estimate of the eigenvalue \( \mu_i \). The quality of the estimate depends upon the size of the off-diagonal elements of \( T \). In fact, the theorem states that if you draw a circle centered at \( T_{ii} \) with radius

\[
r_i = \sum_{j \neq i} |T_{ij}|
\]

i.e., the sum of the absolute values of the off-diagonal elements, then \( \mu_i \) lies somewhere in that circle, as shown in Fig. 13. (For symmetric matrices the eigenvalues are always real, so that they lie on the real axis. Nonsymmetric matrices can have complex eigenvalues, and in such cases the extra dimension implied by the circle is important.) There is a catch. If two circles overlap, then the only thing we can conclude is that both of the two associated eigenvalues lie somewhere in the union of those two circles. For the case illustrated in Fig. 13, we know that \( T_{33} - r_3 \leq \mu_3 \leq T_{33} + r_3 \). We also know that the other two
Fundamentals of Structural Mechanics

Figure 13 Graphical representation of Gershgorin's theorem

eigenvalues satisfy $T_{11} - r_1 \leq \mu_1, \mu_2 \leq T_{22} + r_2$, i.e., they lie somewhere between extremes of the two circles. Clearly, if the off-diagonal elements of the tensor are small, the diagonal elements are very good estimates of the eigenvalues. In any case, the diagonal elements should be good starting points for the Newton iteration. It also provides a means of checking our eigenvalues once we have found them. If they do not lie within the proper bounds, they cannot be correct. This theorem applies to matrices of any dimension.

Once one root is determined, one can use synthetic division to factor the root out of the cubic, leaving a quadratic that can be solved by the quadratic formula. Alternatively, we could simply use Eqn. (53) from another starting point in the hope that it would converge to one of the other roots (there is no guarantee that the iteration will converge to a root different from one already found).

**Determination of the eigenvectors.** The cubic equation has three roots, which we call $\mu_1, \mu_2, \text{ and } \mu_3$. Each of these roots corresponds to an eigenvector. Let the eigenvectors corresponding to $\mu_1, \mu_2, \text{ and } \mu_3$ be called $n_1, n_2,$ and $n_3$, respectively. These eigenvectors can be determined by solving the system of equations $[T - \mu I]n_i = 0$ (no implied sum on $i$). However, by the very definition of the eigenvalues, the coefficient matrix $[T - \mu I]$ is singular, so we must exercise some care in solving these equations.

Let us try to find the eigenvector $n_i$ associated with $\mu_i$ (any one of the eigenvalues). Let us assume that the eigenvector has the form

$$n_i = n_1^{(i)}e_1 + n_2^{(i)}e_2 + n_3^{(i)}e_3$$

Our aim is to determine the, as yet unknown, values of $n_1^{(i)}, n_2^{(i)}, \text{ and } n_3^{(i)}$. To aid the discussion let us define three vectors that have components equal to the columns of the coefficient matrix $[T - \mu_i I]$

$$t_1^{(i)} \sim \begin{bmatrix} T_{11} - \mu_i \\ T_{21} \\ T_{31} \end{bmatrix} \quad t_2^{(i)} \sim \begin{bmatrix} T_{12} \\ T_{22} - \mu_i \\ T_{32} \end{bmatrix} \quad t_3^{(i)} \sim \begin{bmatrix} T_{13} \\ T_{23} \\ T_{33} - \mu_i \end{bmatrix}$$

The equation $[T - \mu_i I]n_i = 0$ can be written as (dropping the superscript "(i)" just to simplify the notation)
Chapter 1  Vectors and Tensors

\[ n_1 t_1 + n_2 t_2 + n_3 t_3 = 0 \]  
\( (55) \)

It should first be obvious that the vectors \( \{ t_1, t_2, t_3 \} \) are not linearly independent. In fact, we selected \( \mu_i \) precisely to create this linear dependence. Besides, if these vectors were linearly independent then, by a theorem of linear algebra, the only possible solution to Eqn. (55) would be \( n_1 = n_2 = n_3 = 0 \), which is clearly at odds with our original aim.

Consider the case where the eigenvalue \( \mu_i \) is distinct (i.e., neither of the other two eigenvalues is equal to it). In this case at least two of the three vectors \( \{ t_1, t_2, t_3 \} \) are linearly independent. The trouble is we do not know in advance which two. There are three possibilities: \( \{ t_1, t_2 \} \), \( \{ t_1, t_3 \} \), and \( \{ t_2, t_3 \} \). We can write Eqn. (55) as

\[ n_\alpha t_\alpha + n_\beta t_\beta = -n_\gamma t_\gamma \]  
\( (56) \)

where no summation is implied and the integers \( \{ \alpha, \beta, \gamma \} \) take on distinct values of 1, 2, or 3 (i.e., no two can be the same). Our three choices are then \( \{ \alpha, \beta, \gamma \} = \{ 1, 2, 3 \} \), \( \{ 2, 3, 1 \} \), or \( \{ 3, 1, 2 \} \). Equation (56) is overdetermined. There are more equations (3) than unknowns (2). However, by construction these equations should be consistent with each other. Hence, any two of the equations should be sufficient to determine \( n_\alpha \) and \( n_\beta \). To remove the ambiguity we can replace Eqn. (56) with its normal form by taking the dot product first with respect to \( t_\alpha \) and then with respect to \( t_\beta \) to give two equations in two unknowns:

\[ \begin{bmatrix} t_\alpha \cdot t_\alpha & t_\alpha \cdot t_\beta \\ t_\beta \cdot t_\alpha & t_\beta \cdot t_\beta \end{bmatrix} \begin{bmatrix} n_\alpha \\ n_\beta \end{bmatrix} = -n_\gamma \begin{bmatrix} t_\alpha \cdot t_\gamma \\ t_\beta \cdot t_\gamma \end{bmatrix} \]  
\( (57) \)

Among the three choices of \( \{ \alpha, \beta, \gamma \} \) at least one must work. Equation (57) will not be solvable if the coefficient matrix is singular. That would be true if its determinant was zero, i.e., if \( (t_\alpha \cdot t_\alpha)(t_\beta \cdot t_\beta) = (t_\alpha \cdot t_\beta)^2 \). If this is the case then it is also true that \( n_\gamma = 0 \), which can certainly be verified once you have successfully solved the problem. If your first choice of \( \{ \alpha, \beta, \gamma \} \) did not work out, then try another one.

One of the important things to notice from Eqn. (57) is that \( n_\alpha \) and \( n_\beta \) can only be determined up to an arbitrary multiplier \( n_\gamma \). To solve the equations one can simply specify a value of \( n_\gamma \) (\( n_\gamma = 1 \) will work just fine). The vector can be scaled by a constant \( Q \) to give the final vector \( n = Q \left( n_\alpha e_\alpha + n_\beta e_\beta + n_\gamma e_\gamma \right) \). The condition of unit length of \( n \) establishes the value of \( Q \) as

\[ Q = \left( n_\alpha^2 + n_\beta^2 + n_\gamma^2 \right)^{-1/2} \]  
\( (58) \)

**Orthogonality of the eigenvectors.** One interesting feature of the eigenvalue problem is that the eigenvectors for distinct eigenvalues are orthogonal, as suggested in the following lemma.
Lemma. Let $n_i$ and $n_j$ be eigenvectors of the symmetric tensor $T$ corresponding to distinct eigenvalues $\mu_i$ and $\mu_j$, respectively (that is, they satisfy $Tn = \mu n$). Then $n_i$ is orthogonal to $n_j$, i.e., $n_i \cdot n_j = 0$.

Proof. The proof is based on taking the difference of the products of the eigenvectors with $T$ in different orders (no summation on repeated indices)

$$0 = n_j \cdot Tn_i - n_i \cdot Tn_j$$

$$= n_j \cdot (\mu_i n_i) - n_i \cdot (\mu_j n_j)$$

$$= (\mu_i - \mu_j) n_i \cdot n_j$$

The first line of the proof is true by definition of symmetry of $T$. The second line substitutes the eigenvalue property $Tn = \mu n$. The last line reflects that the dot product of vectors is commutative. Since we assumed that the eigenvalues were distinct, Eqn. (59) can be true only if $n_i \cdot n_j = 0$, that is, if they are orthogonal. Q

Notice that orthogonality does not hold if the eigenvalues are repeated because Eqn. (59) is satisfied even if $n_i \cdot n_j \neq 0$. We will see the ramification of this observation in the following examination of the special cases.

Special cases. There are two special cases that deserve mention. Both correspond to repeated roots of the characteristic equation. The main concern is how to find the eigenvectors associated with repeated roots.

If $\mu_\alpha = \mu_\beta \neq \mu_\gamma$, we have the case that two of the roots are equal, but the third is distinct. For the distinct root $\mu_\gamma$, we can follow the above procedure and find the unique eigenvector $n_\gamma$. The vectors corresponding to the double eigenvalue are not unique. If we have two eigenvectors $n_\alpha$ and $n_\beta$ corresponding to $\mu_\alpha = \mu_\beta = \mu$, then any vector that is a linear combination of those two vectors, $n = an_\alpha + bn_\beta$, is also an eigenvector. The proof is simple

$$Tn = T(an_\alpha + bn_\beta) = a\mu n_\alpha + b\mu n_\beta = \mu(aw_\alpha + bw_\beta) = \mu n$$

Since the eigenvectors are orthogonal for distinct eigenvalues, the physical interpretation of an eigenvector $n$ corresponding to the double eigenvalue $\mu$ is that it is any vector that lies in the plane normal to $n_\alpha$, as shown in Fig. 14.

There is a clever way of finding such a vector. The tensor $[I - n \otimes n]$ is a projection tensor. When applied to any vector $m$, it will produce a new vector that is orthogonal to $n$. Specifically
Chapter 1  Vectors and Tensors

Figure 14  Physical interpretation of eigenvectors for repeated eigenvalues

\[ \mathbf{m} = (\mathbf{I} - \mathbf{n} \otimes \mathbf{n}) \mathbf{m} = \mathbf{m} - (\mathbf{n} \cdot \mathbf{m}) \mathbf{n} \]  \hspace{1cm} (60)

is orthogonal to \( \mathbf{n} \) (prove it by computing the value of the dot product of vectors \( \mathbf{n} \) and \( \mathbf{m} \)). Thus, to compute the eigenvectors corresponding to the double root, we need only take any vector \( \mathbf{m} \) in the space (not collinear with \( \mathbf{n}_x \)) and compute

\[ \mathbf{n}_p = \mathbf{m} - (\mathbf{n}_y \cdot \mathbf{m}) \mathbf{n}_y \]  \hspace{1cm} (61)

then normalize as \( \mathbf{n}_p = \mathbf{n}_p/\| \mathbf{n}_p \| \). To get a third eigenvector that is orthogonal to the other two, we can simply compute the cross product \( \mathbf{n}_a = \mathbf{n}_p \times \mathbf{n}_y \).

The second special case has all three of the eigenvalues equal, \( \mu_1 = \mu_2 = \mu_3 = \mu \). In this case, any vector in the space is an eigenvector. If we need an orthonormal set of three specific vectors, we can apply the same procedure as before, starting with any two (noncollinear) vectors.

**Example 3. Distinct roots.** Consider that the components of the tensor \( \mathbf{T} \) are given by the matrix of values

\[
\mathbf{T} = \begin{bmatrix}
3 & -1 & 0 \\
-1 & 3 & 0 \\
0 & 0 & 3
\end{bmatrix}
\]

The invariants are \( I_x = 9, \ II_x = 26, \) and \( III_x = 24 \). The characteristic equation for the eigenvalues is \( -\mu^3 + 9\mu^2 - 26\mu + 24 = 0 \). This equation can be factored (not many real problems have integer roots!) as

\[-(\mu - 2)(\mu - 3)(\mu - 4) = 0\]

showing that the roots are \( \mu_1 = 2, \mu_2 = 3, \) and \( \mu_3 = 4 \). (Note that Gershgorin’s theorem holds!) The eigenvector associated with the first eigenvalue can be found by solving the equation \( [\mathbf{T} - \mu_i \mathbf{I}] \mathbf{n}_i = 0 \). We can observe that

\[
[\mathbf{T} - \mu_i \mathbf{I}] = \begin{bmatrix}
1 & -1 & 0 \\
-1 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix} \Rightarrow \begin{aligned}
\mathbf{t}_1^{(1)} &= \mathbf{e}_1 - \mathbf{e}_2 \\
\mathbf{t}_2^{(1)} &= -\mathbf{e}_1 + \mathbf{e}_2 \\
\mathbf{t}_3^{(1)} &= \mathbf{e}_3
\end{aligned}
\]

Taking the choice \( \{\alpha, \beta, \gamma\} = \{2, 3, 1\} \), Eqn. (56) gives

\[ n_2^{(1)} t_2^{(1)} + n_3^{(1)} t_3^{(1)} = -n_1^{(1)} t_1^{(1)} \]
Letting \( n_1^{(1)} = 1 \), the normal equations, Eqn. (57), take the form

\[
\begin{bmatrix}
2 & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
2 \\
0
\end{bmatrix}
\]

which gives \( n_2^{(1)} = 1 \) and \( n_3^{(1)} = 0 \). Thus, the eigenvector for \( \mu_1 = 2 \) is

\[
n_1 = e_1 + e_2
\]

The remaining two eigenvectors can be found in exactly the same way, and are

\[
n_2 = e_3, \quad n_3 = e_1 - e_2
\]

These vectors can, of course, be normalized to unit length.

It is interesting to note what happens for other choices of the normal equations in the preceding example. In particular, it is evident that \( t_2^{(1)} = t_3^{(1)} \). If we were to make the choice \( \{a, \beta, \gamma\} = \{1, 2, 3\} \) then the coefficient matrix for the normal equations would be singular. This observation is also consistent with the fact that \( n_3^{(1)} = 0 \).

**Example 4. Repeated roots.** Consider that the components of the tensor \( T \) are given by the matrix of values

\[
T = \begin{bmatrix}
5 & -1 & -1 \\
-1 & 5 & -1 \\
-1 & -1 & 5
\end{bmatrix}
\]

The invariants are \( I_T = 15 \), \( II_T = 72 \), and \( III_T = 108 \). The characteristic equation for the eigenvalues is

\[
-\mu^3 + 15\mu^2 - 72\mu + 108 = 0
\]

or

\[
-(\mu - 3)(\mu - 6)(\mu - 6) = 0
\]

showing that the roots are \( \mu_1 = \mu_2 = 6 \), and \( \mu_3 = 3 \). The eigenvector associated with the distinct eigenvalue \( \mu_3 \) can be found by solving the equation \( [T - \mu_3 I]n_3 = 0 \) as in the previous example. The result is

\[
n_3 = \frac{1}{\sqrt{3}}(e_1 + e_2 + e_3)
\]

The eigenvectors corresponding to the repeated root must lie in a plane orthogonal to \( n_3 \). We can select any vector in the space and project out the component along \( n_3 \). Let us use \( m = e_1 \). Project out the part of the vector along \( n_3 \) (see Example 1)
Chapter 1  Vectors and Tensors

\[ n_2 = \varrho P m = \varrho \left[ 1 - n_3 \otimes n_3 \right] e_1 \]
\[ = \varrho \left[ e_1 - (n_3 \cdot e_1) n_3 \right] \]
\[ = \varrho \left[ e_1 - \frac{1}{3} (e_1 + e_2 + e_3) \right] \]
\[ = \varrho \left( \frac{2}{3} e_1 - \frac{1}{3} e_2 - \frac{1}{3} e_3 \right) \]
\[ = \frac{1}{\sqrt{6}} \left( 2e_1 - e_2 - e_3 \right) \]

where the constant \( \varrho \) was selected to give the vector unit length. Finally, \( n_1 \) can be computed as \( n_1 = n_2 \times n_3 \) to give

\[ n_1 = \frac{1}{\sqrt{2}} (-e_2 + e_3) \]

The spectral decomposition. If the eigenvalues and eigenvectors are known, we can express the original tensor in terms of those objects in the following manner

\[
T = \sum_{i=1}^{3} \mu_i n_i \otimes n_i \tag{62}
\]

Note that we need to suspend the summation convention because of the number of times that the index \( i \) appears in the expression. This form of expression of the tensor \( T \) is called the spectral decomposition of the tensor. How do we know that the tensor \( T \) is equivalent to its spectral decomposition? As we indicated earlier, the operation of a second-order tensor is completely defined by its operation on three independent vectors. Let us assume that the eigenvectors \( \{ n_1, n_2, n_3 \} \) are orthogonal (which means that any eigenvectors associated with repeated eigenvalues were orthogonalized). Let us examine how the tensor and its spectral decomposition operate on \( n_j \)

\[ T n_j = \sum_{i=1}^{3} \mu_i \left[ n_i \otimes n_i \right] n_j = \sum_{i=1}^{3} \mu_i \left( n_j \cdot n_i \right) n_i = \sum_{i=1}^{3} \mu_i \delta_{ij} n_i = \mu_j n_j \]

Thus, we have concluded that both tensors operate the same way on the three eigenvectors. Therefore, the spectral representation must be equivalent to the original tensor. A corollary of the preceding construction is that any two tensors with exactly the same eigenvalues and eigenvectors are equivalent.

The spectral decomposition affords us another remarkable observation. We know that we are free to select any basis vectors to describe the components of a tensor. What happens if we select the eigenvectors \( \{ n_1, n_2, n_3 \} \) as the basis? According to Eqn. (62), in this basis the off-diagonal components of the tensor \( T \) are all zero, while the diagonal elements are exactly the eigenvalues
The invariants of \( T \) also take a special form when expressed in terms of the eigenvalues. The invariants are, by their very nature, independent of the basis chosen to represent the tensor. As such, one must get the same value of the invariants in all bases. Those values will, of course, be the values computed in any specific basis. The simplest basis, often referred to as the canonical basis, is the one given by the eigenvectors. In this basis, the invariants can be represented as

\[
I_T = \mu_1 + \mu_2 + \mu_3 \\
II_T = \mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3 \\
III_T = \mu_1\mu_2\mu_3
\]

(63)

**Example 5.** Consider a tensor \( T \) that has one distinct eigenvalue \( \mu_1 \) and a repeated eigenvalue \( \mu_2 = \mu_3 \). Use the spectral decomposition to show that the tensor \( T \) can be represented as

\[
T = \mu_1 \left[ n \otimes n \right] + \mu_2 \left[ I - n \otimes n \right]
\]

where \( n \) is the unit eigenvector associated with the distinct eigenvalue \( \mu_1 \).

Let \( n_1 = n, n_2, \) and \( n_3 \) be eigenvectors of \( T \). Further assume that these vectors are orthogonal (remember, if they are not orthogonal due to a repeated root, they can always be orthogonalized). The sum of outer products of orthonormal vectors is the identity. Thus,

\[
I = \sum_{i=1}^{3} n_i \otimes n_i = n \otimes n + \sum_{i=2}^{3} n_i \otimes n_i
\]

Write \( T \) in terms of its spectral decomposition as

\[
T = \sum_{i=1}^{3} \mu_i \left[ n_i \otimes n_i \right] = \mu_1 n \otimes n + \sum_{i=2}^{3} \mu_i \left[ n_i \otimes n_i \right]
\]

\[
= \mu_1 n \otimes n + \mu_2 \sum_{i=2}^{3} n_i \otimes n_i
\]

\[
= \mu_1 n \otimes n + \mu_2 \left[ I - n \otimes n \right]
\]

There is great significance to this result. Notice that the final spectral representation does not refer to \( n_2 \) and \( n_3 \) at all. Since these vectors are arbitrarily chosen from the plane orthogonal to \( n \) these vectors have no intrinsic significance (other than that they faithfully represent the plane). In this case there are only three in-
trinsic bits of information: $\mu_1$, $\mu_2$, and $n$. Hence, this representation of $T$ is canonical.

The Cayley-Hamilton theorem. The spectral decomposition and the characteristic equation for the eigenvalues of a tensor can be used to prove the Cayley-Hamilton theorem, which states that

$$T^3 - I_T T^2 + H_T T - H_T I = 0 \quad (64)$$

where $T^2 = T T$ and $T^3 = T T T$ are products of the tensor $T$ with itself. Using the spectral decomposition, one can show that (Problem 22)

$$T^n = \sum_{i=1}^{3} (\mu_i)^n n_i \otimes n_i$$

Using this result, and noting that $I = n_i \otimes n_i$ (sum implied), we can compute

$$T^3 - I_T T^2 + H_T T - H_T I = \sum_{i=1}^{3} (\mu_i^3 - I_T \mu_i^2 + H_T \mu_i - H_T)n_i \otimes n_i$$

All of the eigenvalues satisfy the characteristic equation. Thus, the term in parentheses is always zero, thereby proving the theorem.

Vector and Tensor Calculus

A field is a function of position defined on a particular region. In our study of mechanics we shall have need of scalar, vector, and tensor fields, in which the output of the function is a scalar, vector, or tensor, respectively. For problems defined on a region of three-dimensional space, the input is the position vector $x$. A function defined on a three-dimensional domain, then, is a function of three independent variables (the components $x_1$, $x_2$, and $x_3$ of the position vector $x$). In certain specialized theories (e.g., beam theory, plate theory, and plane stress) position will be described by one or two independent variables.

A field theory is a physical theory built within the framework of fields. The primary advantage of using field theories to describe physical phenomena is that the tools of differential and integral calculus are available to carry out the analysis. For example, we can appeal to concepts like infinitesimal neighborhoods and limits. And we can compute rates of change by differentiation and accumulations and averages by integration.

Figure 15 shows the simplest possible manifestation of a field: a scalar function of a scalar variable, $g(x)$. A scalar field can, of course, be represented as a graph with $x$ as the abscissa and $g(x)$ as the ordinate. For each value of position $x$ the function produces as output $g(x)$. The derivative of the function is defined through the limiting process as
A scalar field \( g(x) \) is a scalar-valued function of a scalar variable. Differentiation gives the slope of the curve at a point and integration gives the area \( A \) under the curve between two points.

\[
\frac{dg}{dx} = \lim_{\Delta x \to 0} \left( \frac{g(x + \Delta x) - g(x)}{\Delta x} \right) = g'(x) \tag{65}
\]

The derivative has the familiar geometrical interpretation of the slope of the curve at a point and gives the rate of change of \( g \) with respect to change in position \( x \). Many of the graphical constructs that serve so well for scalar functions of scalar variables do not generalize well to vector and tensor fields. However, the concept of the derivative as the limit of the ratio of flux, \( g(x + \Delta x) - g(x) \) in the present case, to size of the region, \( \Delta x \) in the present case, will generalize for all cases.

Figure 16 illustrates that a segment \( [x, x + \Delta x] \) has a left end and a right end. If we ascribe a directionality to the segment by imagining the positive direction to be in the direction of the +\( x \) axis, then the left end is the "inflow" boundary and the right end is the "outflow" boundary of the segment. We can think of the flux of \( g \) as being the difference between the outflow and the inflow. For a scalar function of a scalar variable that is simply \( g(x + \Delta x) - g(x) \). According to Eqn. (65), the derivative \( \frac{dg}{dx} \) is the limit of the ratio of flux to size of the region.

In three-dimensional space we shall generalize our concept of derivative (rate of change) using an arbitrary region \( \Omega \) having volume \( V(\Omega) \) with surface \( \Omega \) having unit normal vector field \( \mathbf{n} \), as shown in Fig. 17. We will define various

\[
\frac{dg}{dx} = \lim_{\Delta x \to 0} \left( \frac{g(x + \Delta x) - g(x)}{\Delta x} \right) = g'(x) \tag{65}
\]
types of derivatives of various types of fields in the following sections, but all of these derivatives will be the limit of the ratio of some sort of flux (outflow minus inflow) to the volume of the region as the volume shrinks to zero. In these definitions the flux will involve an integral over the surface area and the normal vector $\mathbf{n}$ will help to distinguish "inflow" from "outflow" for the situation at hand. For each definition of derivative we will develop a coordinate expression that will tell us how to formally "take the derivative" of the field. The coordinate expressions will all involve partial derivatives of the vector or tensor components.

The integral of the function between the limits $b$ and $c$ gives the area between the graph of the function $g(x)$ and the $x$ axis (see Fig. 15). For any scalar function of a scalar variable one can think of the integral as the "area under the curve." Integration is the limit of a sum of infinitesimal strips with area $g(x) \, dx$. The total area is the accumulated sum of the infinitesimal areas. The geometric notion of integration is quite independent of techniques of integration based upon anti-derivatives of functions because there are methods of integration (e.g., numerical quadrature) that do not rely upon the anti-derivative. In our developments here we need to think of integrals both in the sense of executing integrals (mostly later in the book) and in the more generic sense of accumulating the limit of a sum.

In three dimensional space we will encounter surface integrals and volume integrals. Most of the time we will not use the notation of "double integrals" for surface integrals and "triple integration" for volume integrals, but rather understand that

$$\int \mathbf{\cdot} \, dA = \int \int (\mathbf{\cdot}) \, dx \, dy, \quad \int \mathbf{\cdot} \, dV = \int \int \int (\mathbf{\cdot}) \, dx \, dy \, dz \quad (66)$$

where the variables and infinitesimals must be established for the coordinate system that is being used to characterize the problem at hand. Again, techniques of integration are important only in particular problems to carry out computations.

The second aspect of integration that we will introduce in this chapter is the idea of integral theorems that provide an equivalence between a surface inte-
Fundamentals of Structural Mechanics

Scalar fields of vector variables. A scalar field is a function $g(x)$ that assigns a scalar value to each point $x$ in a particular domain. The temperature in a solid body is an example of a scalar field. As an example consider the scalar field $g(x) = \| x \|^2 = x_1^2 + x_2^2 + x_3^2$, in which the function $g(x)$ gives the square of the length of the position vector $x$. In two dimensions, a scalar field can be represented by either a graph or a contour map like those shown in Fig. 18.

As with any function that varies from point to point in a domain, we can ask the question: At what rate does the field change as we move from one point to another? It is fairly obvious from the contour map that if one moves from one point to another along a contour then the change in the value of the function is zero (and therefore the rate of change is zero). If one crosses contours then the function value changes. Clearly, the question of rate of change depends upon direction of the line connecting the two points in question.

Consider a scalar field $g$ in three dimensional space evaluated at two points $a$ and $b$, as shown in Fig. 19. Point $a$ is located at position $x$ and point $b$ is located at position $x + \Delta s n$, where $n$ is a unit vector that points in the direction from $a$ to $b$ and $\Delta s$ is the distance between them. The directional derivative of the function $g$ in the direction $n$, denoted $Dg \cdot n$, is the ratio of the difference in the function values at $a$ and $b$ to the distance between the points, as the point $b$ is taken closer and closer to $a$.

\[
\int_{a}^{b} \left( \frac{dg}{dx} \right) dx = g(b) - g(a) \quad (67)
\]

The remainder of this chapter is devoted to reviewing of some of the basic ideas from vector calculus and the extension of those ideas to tensor fields.

Figure 18 (a) A graph and (b) a contour map of a scalar field in two dimensions.
The directional derivative of \( g \) can be computed, using the chain rule of differentiation, from the formula

\[
Dg(x) \cdot n = \lim_{\Delta s \to 0} \frac{g(x + \Delta s n) - g(x)}{\Delta s}
\]

The directional derivative of \( g \) can be computed, using the chain rule of differentiation, from the formula

\[
Dg(x) \cdot n = \frac{d}{de} \left( g(x + \varepsilon n) \right)_{\varepsilon=0} = \frac{\partial g}{\partial x_i} n_i
\]

In essence, the directional derivative determines the one-dimensional rate of change (i.e., \( d/de \)) of the function at the point \( x \) and just starting to move in the fixed direction \( n \). Because \( x \) and \( n \) are fixed, the derivative is an ordinary one.

**Example 6. Directional Derivative.** Consider the scalar function given by the expression \( g(x) \equiv x \cdot x = x_k x_k \). We can compute the directional derivative in the direction \( n \) by Eqn. (69). Noting that the augmented function can be written as \( g(x + \varepsilon n) = (x_k + \varepsilon n_k)(x_k + \varepsilon n_k) \), we compute the directional derivative as

\[
Dg(x) \cdot n = \frac{d}{de} \left[ (x_k + \varepsilon n_k)(x_k + \varepsilon n_k) \right]_{\varepsilon=0}
\]

\[
= \frac{d}{de} [x_k x_k + 2 \varepsilon x_k n_k + \varepsilon^2 n_k n_k]_{\varepsilon=0}
\]

\[
= [2x_k n_k + 2\varepsilon n_k n_k]_{\varepsilon=0} = 2x_k n_k
\]

It is also useful to note that \( \frac{\partial g}{\partial x_i} = \delta_{ik} x_k + x_k \delta_{ki} = 2x_k \). Then, according to Eqn. (69) again, we have

\[
Dg \cdot n = \frac{\partial g}{\partial x_i} n_i = 2x_i n_i
\]

which is identical to the previous result.

From Eqn. (69) it is evident that the partial derivatives of the function \( g \) play a key role in determining the rate of change in a particular direction. In fact, the partial derivatives \( \frac{\partial g}{\partial x_i} \) give the rate of change of \( g \) in the direction of the coordinate axis \( x_i \). These three quantities can be viewed as the components of
a vector called the gradient of the field. The gradient of a scalar field \( g(x) \) is a vector field \( \nabla g(x) \), which, in Cartesian coordinates, is given by

\[
\nabla g(x) = \frac{\partial g(x)}{\partial x_j} e_j
\]  

(70)

where summation on \( j \) is implied. With this definition of the gradient, the directional derivative takes the form

\[
Dg \cdot n = \nabla g \cdot n
\]

(71)

We know that the directional derivative of \( g \) is zero if \( n \) is tangent to a contour line. Therefore, the vector \( \nabla g \) must be perpendicular to the contour lines, as shown in Fig. 18b, because \( \nabla g \cdot n = 0 \) in that direction. For the direction \( n = \nabla g/|| \nabla g || \) it is evident from Eqn. (71) that \( Dg \cdot n = || \nabla g(x) ||. \) Hence, \( || \nabla g(x) || \) is the maximum rate of change of the scalar field \( g \).

We can define the gradient of a scalar field independent of any coordinate system. Consider an arbitrary region \( \Omega \) with surface \( \Omega \) and outward unit normal vector field \( n \), shown in Fig. 17. The gradient is the ratio of the flux \( gn \) over the surface to the volume \( \mathcal{V}(\Omega) \), in the limit as the volume of the region shrinks to zero. To wit

\[
\nabla g = \lim_{\mathcal{V}(\Omega) \to 0} \frac{1}{\mathcal{V}(\Omega)} \int_\Omega g \ n \ dA
\]

(72)

where \( \mathcal{V}(\Omega) \) is the volume of the enclosed surface.

Equation (72) does not depend upon a specific coordinate system. Equation (70) is a formula for the gradient in rectangular Cartesian coordinates. The derivation of Eqn. (70) from Eqn. (72) is very instructive. To compute with Eqn. (72) we need to select a specific region \( \Omega \) so that we can compute the flux and the volume and take the limit as the volume shrinks to zero. The simplest possible choice is the cuboid with sides parallel to the coordinate planes shown in Fig. 20. The volume of this region is \( \mathcal{V}(\Omega) = \Delta x_1 \Delta x_2 \Delta x_3 \). The surface \( \Omega \) consists of six rectangles each with constant normal \( n \) pointing in the direction of one of the base vectors. Furthermore, the six faces occur in pairs with normals \( n = \pm e_i \) on which \( x_i \) is constant over the entire face (with a value of \( x_i \) for the face with normal \( -e_i \) and \( x_i + \Delta x_i \) for the face with normal \( e_i \)). Hence, we can compute the flux as

\[
\int_\Omega g \ n \ dA = \sum_{i=1}^3 \left[ g(x + \Delta x_i e_i) e_i + g(x)(-e_i) \right] dA_i
\]

(73)

where \( \Omega_i \) is the rectangular region with area \( A_i \) over which \( x_i \) is constant. Note that \( A_1 = \Delta x_2 \Delta x_3, A_2 = \Delta x_3 \Delta x_1, \) and \( A_3 = \Delta x_1 \Delta x_2 \) are the areas of the
Figure 20 A particular region for the computation of flux and volume needed to compute derivatives in multivariate calculus

faces. Next, we can recognize that the volume is \( V(B) = A_i \Delta x_i \) (no sum) for any \( i=1, 2, 3 \). Finally, we can recognize that

\[
\frac{1}{A_i} \int_{\Omega_i} (\cdot) \, dA_i
\]

is simply the average of \((\cdot)\) over the integration region \( \Omega_i \). In the limit, as the volume and the face areas shrink to zero, the average values will approach the values at \( x \). Therefore, Eqn. (72) can be written as

\[
\nabla g = \sum_{i=1}^{3} \lim_{\Delta x_i \to 0} \frac{1}{A_i} \int_{\Omega_i} \lim_{\Delta x_i \to 0} \left( \frac{g(x + \Delta x_i e_i) - g(x)}{\Delta x_i} \right) \, dA_i \, e_i
\]

\[
= \sum_{i=1}^{3} \frac{\partial g(x)}{\partial x_i} \, e_i
\]

(75)

The limiting process for \( \Delta x_i \) can be moved inside the integral over \( \Omega_i \) because \( x_i \) is constant for that integral. This limit is, of course, the partial derivative of \( g \) with respect to \( x_i \). That partial derivative is a function of the other two variables which are not constant over that face. However, we then take the limit of the average over the region of integration to give the final result.

As we shall see, this approach will work in essentially identical fashion for developing coordinate expressions for all of the derivatives in this chapter.

**Vector fields.** A vector field is a function \( \mathbf{v}(x) \) that assigns a vector to each point \( x \) in a particular domain. The displacement of a body is a vector field. Each point of the body moves by some amount in some direction. The force induced by gravitational attraction is a vector field.

Figure 21 shows two examples of vector fields. The pictures show the vectors at only enough points to get the idea of how the vectors are oriented and sized. The second vector field shown in the figure can be expressed in functional form as

\[
\mathbf{v}(x) = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2
\]

(76)
The vectors point in the radial direction, and their length is equal to the distance of the point of action to the origin.

In general, if our base vectors are assumed to be constant throughout our domain, then the vector field can be expressed in terms of component functions

\[ \mathbf{v}(\mathbf{x}) = v_i(\mathbf{x}) \mathbf{e}_i \]  

(77)

For example, from Eqn. (76) we can see that the explicit expression for the components of the vector field are \( v_1(\mathbf{x}) = x_1, v_2(\mathbf{x}) = x_2, \) and \( v_3(\mathbf{x}) = 0. \) For curvilinear coordinates, the base vectors are also functions of the coordinates.

There are as many ways to differentiate a vector field as there are ways of multiplying vectors. The analogy between vector multiplication and vector differentiation is given in the following table

<table>
<thead>
<tr>
<th>Multiplication</th>
<th>Differentiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{u} \cdot \mathbf{v} ) dot</td>
<td>( \text{div} , \mathbf{v} ) divergence</td>
</tr>
<tr>
<td>( \mathbf{u} \times \mathbf{v} ) cross</td>
<td>( \text{curl} , \mathbf{v} ) curl</td>
</tr>
<tr>
<td>( \mathbf{u} \otimes \mathbf{v} ) tensor</td>
<td>( \nabla \mathbf{v} ) gradient</td>
</tr>
</tbody>
</table>

As was the case for vector multiplication, each different way to differentiate a vector field yields a result with different character. For example, the divergence of a vector field is a scalar field, while the gradient of a vector field is a tensor field. Each of these derivatives, however, represents the rate of change of the vector field in some sense. Each one can be viewed as the "first derivative" of the vector field. In the sequel, we shall give a definition for each of these derivatives and give an idea of what they physically represent.

**The divergence of a vector field.** One way to measure of the rate of change of a vector field is the divergence. Consider again a domain \( \mathcal{B} \) with enclosed volume \( \mathcal{V}(\mathcal{B}) \) and boundary \( \Omega \) with unit normal vector \( \mathbf{n} \), as shown in Fig. 17. Let us assume that the body lives in a vector field \( \mathbf{v}(\mathbf{x}) \). Thus, at each point \( \mathbf{x} \) in \( \mathcal{B} \) there exists a vector \( \mathbf{v}(\mathbf{x}) \). Let the flux be \( \mathbf{v} \cdot \mathbf{n} \) on the boundary \( \Omega \). The

![Figure 21](image-url)  

*Figure 21* A vector field assigns a vector to each point in a domain
The divergence of the vector field is defined as the limit of the ratio of flux to volume, in the limit as the volume shrinks to zero. To wit

$$\text{div}(\mathbf{v}) \equiv \lim_{V(\mathcal{B}) \to 0} \frac{1}{V(\mathcal{B})} \int_{\Omega} \mathbf{v} \cdot \mathbf{n} \, dA$$

where \( dA \) is the infinitesimal element of area defined on the surface.

We can better understand why the integrand \( \mathbf{v} \cdot \mathbf{n} \) is called the flux if we think of the vector field \( \mathbf{v} \) as the particle velocity in a fluid flow, wherein the vectors would be tangent to particle streamlines. The product \( \mathbf{v} \cdot \mathbf{n} \) would then represent the total amount of fluid that escapes through the area \( dA \) on the boundary per unit of time, as shown in Fig. 22. The physical significance of the product \( \mathbf{v} \cdot \mathbf{n} \) is that the volume of fluid that passes through the area \( dA \) in unit time is equal to the base area of the cylinder \( dA \) times the height of the cylinder \( \mathbf{v} \cdot \mathbf{n} \). Note that streamlines that are tangent to the boundary (i.e., \( \mathbf{v} \cdot \mathbf{n} = 0 \)) do not let any fluid out, while streamlines normal to the boundary let it out most efficiently.

Let us compute an expression for the divergence of a vector field in Cartesian coordinates, again using the simple cuboid shown in Fig. 20. Following the same conventions we can compute the flux as

$$\int_{\Omega} \mathbf{v} \cdot \mathbf{n} \, dA = \sum_{i=1}^{3} \int_{\Omega_i} \left[ \mathbf{v}(x + \Delta x_i \mathbf{e}_i) \cdot \mathbf{e}_i + \mathbf{v}(x) \cdot (-\mathbf{e}_i) \right] dA_i$$

where, again, \( \Omega_i \) is the rectangular region with area \( A_i \) over which \( x_i \) is constant. Substituting \( A_i = V(\mathcal{B})/\Delta x_i \), we get

$$\text{div}(\mathbf{v}) = \sum_{i=1}^{3} \lim_{\Delta x_i \to 0} \frac{1}{A_i} \int_{\Omega_i} \lim_{\Delta x_i \to 0} \left( \frac{\mathbf{v}(x + \Delta x_i \mathbf{e}_i) - \mathbf{v}(x)}{\Delta x_i} \right) dA_i \cdot \mathbf{e}_i$$

Taking the limit of the average of the limit, as before, we arrive at the expression for the divergence in Cartesian coordinates:

$$\text{div}(\mathbf{v}) = \frac{\partial \mathbf{v}(x)}{\partial x_i} \cdot \mathbf{e}_i = \frac{\partial v_i(x)}{\partial x_i}$$
Note that the summation convention applies to indices that are repeated in a quotient. A common notation for the partial derivative is \( (\cdot)_i \equiv \partial (\cdot)/\partial x_i \). This notation is usually referred to as the *comma notation* for partial derivatives. This notation is useful if there is no ambiguity the variable of differentiation. In this abbreviated notation, the divergence has the more compact expression \( \text{div}(\mathbf{v}) = v_{,i} \) with summation implied across the comma. It should be evident that the comma notation is convenient for index manipulation.

**The gradient of a vector field.** Consider again the domain \( \Omega \) with boundary \( \partial \Omega \) shown in Fig. 17. The gradient of a vector field \( \mathbf{v}(\mathbf{x}) \) is a second-order tensor defined as the limit of the ratio of the flux \( \mathbf{v} \otimes \mathbf{n} \) over the surface to the volume, as the volume shrinks to zero. To wit

\[
\nabla \mathbf{v} \equiv \lim_{\mathcal{V}(\Omega) \to 0} \frac{1}{\mathcal{V}(\Omega)} \int_{\partial \mathcal{V}} \mathbf{v} \otimes \mathbf{n} \, dA
\]

Again, \( \mathcal{V}(\Omega) \) is the volume of the region \( \Omega \), \( \partial \mathcal{V} \) is the surface of the region, and \( \mathbf{n} \) is the unit normal vector field to the surface. With a construction similar to the one used for the divergence, we can compute a coordinate expression for the gradient. The component expression for \( \nabla \mathbf{v} \) in Cartesian coordinates is

\[
\nabla \mathbf{v} = \frac{\partial v_i(\mathbf{x})}{\partial x_j} \left[ \mathbf{e}_i \otimes \mathbf{e}_j \right]
\]

where summation is implied for both \( i \) and \( j \). Thus, the components of \( \nabla \mathbf{v} \) are simply the various partial derivatives of the component functions with respect to the coordinates, that is, the component \( [\nabla \mathbf{v}]_{ij} \) gives the rate of change of the \( i \)th component of \( \mathbf{v} \) with respect to the \( j \)th coordinate axis.

We can interpret the gradient of a vector field geometrically by considering the construction shown in Fig. 23. Consider two points \( a \) and \( b \) that are near to each other (i.e., \( \Delta \mathbf{s} \) is very small). The unit vector \( \mathbf{n} \) points in the direction from \( a \) to \( b \). The value of the vector field at \( a \) is \( \mathbf{v}(\mathbf{x}) \) and the value of the vector field

![Figure 23](image-url)
at \( b \) is \( \mathbf{v}(x + \Delta s \mathbf{n}) \). Since the vector field changes with position in the domain, these two vectors are different in both length and orientation. If we transport a copy of \( \mathbf{v}(x) \) and position it at \( b \) (shown dotted), then we can compare the difference between the two vectors. The vector that connects the head of \( \mathbf{v}(x) \) to the head of \( \mathbf{v}(x + \Delta s \mathbf{n}) \) is \( \mathbf{v}(x + \Delta s \mathbf{n}) - \mathbf{v}(x) \). This vector represents the difference in the vector between points \( a \) and \( b \). If we divide this difference by \( \Delta s \), then we get the rate of change as we move in the specified direction. Finally, taking the limit as \( \Delta s \) goes to zero, we get the directional derivative

\[
D\mathbf{v}(x) \cdot \mathbf{n} = \lim_{\Delta s \to 0} \frac{\mathbf{v}(x + \Delta s \mathbf{n}) - \mathbf{v}(x)}{\Delta s}
\]

Like the analogous formula for scalar fields, the quantity \( D\mathbf{v}(x) \cdot \mathbf{n} \) is called the directional derivative because it gives the rate of change of the vector field in the direction \( \mathbf{n} \). The limiting process above suggests that we can compute the directional derivative as

\[
D\mathbf{v}(x) \cdot \mathbf{n} = \frac{d}{d\varepsilon} \left[ \mathbf{v}(x + \varepsilon \mathbf{n}) \right]_{\varepsilon = 0}
\] (84)

A straightforward application of the chain rule for differentiation gives

\[
D\mathbf{v}(x) \cdot \mathbf{n} = [\nabla \mathbf{v}] \mathbf{n}
\] (85)

The directional derivative provides the answer to the question: What is the rate of change of the vector field? But Eqn. (85) makes it clear that the tensor \( \nabla \mathbf{v} \) contains all of the information needed to assess rate of change in any direction.

**Example 7.** Consider a vector field given by the following explicit expression

\[
\mathbf{v}(x) = x_1 x_2 x_3 (x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3).
\]

The components of the vector field are given by the following expressions

\[

\begin{align*}
\nu_1 &= x_1^2 x_2 x_3, \\
\nu_2 &= x_1 x_2^2 x_3, \\
\nu_3 &= x_1 x_2 x_3^2
\end{align*}
\]

The gradient of this vector field can be computed from Eqn. (83). The result is the following tensor field

\[
\nabla \mathbf{v}(x) = 2x_1 x_2 x_3 [\mathbf{e}_1 \otimes \mathbf{e}_1] + x_1^2 x_3 [\mathbf{e}_1 \otimes \mathbf{e}_2] + x_1^2 x_2 [\mathbf{e}_1 \otimes \mathbf{e}_3] + x_1 x_2^2 [\mathbf{e}_2 \otimes \mathbf{e}_1] + 2x_1 x_2 x_3 [\mathbf{e}_2 \otimes \mathbf{e}_2] + x_1 x_3^2 [\mathbf{e}_2 \otimes \mathbf{e}_3] + x_2 x_3^2 [\mathbf{e}_3 \otimes \mathbf{e}_1] + x_1 x_3^2 [\mathbf{e}_3 \otimes \mathbf{e}_2] + 2x_1 x_2 x_3 [\mathbf{e}_3 \otimes \mathbf{e}_3]
\]

The components of the tensor \( \nabla \mathbf{v} \) can be put in matrix form as follows

\[
\nabla \mathbf{v} \sim \begin{bmatrix}
2x_1 x_2 x_3 & x_1^2 x_3 & x_1^2 x_2 \\
x_1 x_2^2 x_3 & 2x_1 x_2 x_3 & x_1 x_3^2 \\
x_2 x_3^2 & x_1 x_3^2 & 2x_1 x_2 x_3
\end{bmatrix}
\]
The divergence of a vector field can be computed from Eqn. (81). It is worth noting that the divergence is simply the trace of the gradient

$$\text{div}(\mathbf{v}) = \text{tr}(\nabla \mathbf{v})$$

where the trace is the sum of the diagonal components of the tensor. Therefore, for the present example, \(\text{div}(\mathbf{v}) = 6x_1x_2x_3\).

One can define the curl of a vector field in a completely analogous way by considering the flux \(\mathbf{v} \times \mathbf{n}\) (see Problem 45). The details are left to the reader.

A comment on notation for derivatives. There are many notations used to characterize operation in vector calculus. In this book we stick to "div" and \(\nabla\) (some authors use "grad"). Occasionally it is useful to use a shorthand notation for gradients of scalar and vectors fields

$$\nabla g = \frac{\partial g}{\partial x}, \quad \nabla \mathbf{v} = \frac{\partial \mathbf{v}}{\partial x}$$

(86)

While this notation is a bit sloppy it is convenient. For many problems in mechanics we use more than one coordinate system. When we take derivatives we must specify the variable of differentiation (if it is ambiguous). For the divergence we will often use "div" and "DIV" to distinguish between two choices. For the gradient we will often use the notation \(\nabla_s(\cdot)\) or \(\nabla_s(\cdot)\) to indicate the variable of differentiation.

Divergence of a tensor field. A tensor field is a function that assigns a tensor \(\mathbf{T}(\mathbf{x})\) to each point \(\mathbf{x}\) in the domain. Consider a tensor field \(\mathbf{T}(\mathbf{x})\) on a region \(\mathcal{B}\) with surface \(\Omega\) having unit normal vector field \(\mathbf{n}\). There are many ways to differentiate a tensor field. In solid mechanics we are primarily interested in one way. By analogy with vector differentiation, we define the divergence of a tensor field

$$\text{div}\mathbf{T} \equiv \lim_{r(\mathcal{B}) \to 0} \frac{1}{\mathcal{V}(\mathcal{B})} \int_{\Omega} \mathbf{T} \mathbf{n} \, dA$$

(87)

where, as before, \(\mathcal{V}(\mathcal{B})\) is the volume of the region \(\mathcal{B}\), \(\Omega\) is the surface of the region, and \(\mathbf{n}\) is the unit normal vector field to the surface. Since the integrand \(\mathbf{T}\mathbf{n}\) is a vector, \(\text{div}\mathbf{T}\) is a vector.

One can use the definition of the divergence to compute a component expression and to prove the divergence theorem for tensor fields, by following the same arguments we have used for vector fields. Let us compute an expression for the divergence of a tensor field in Cartesian coordinates, again using the simple cuboid shown in Fig. 20. Following the same conventions we can compute the flux as
\begin{equation}
\int_{\Omega} T n \, dA = \sum_{i=1}^{3} \int_{\Omega_i} \left[ T(x + \Delta x_i e_i) e_i + T(x)(-e_i) \right] dA_i
\end{equation}

where, again, $\Omega_i$ is the rectangular region with area $A_i$ over which $x_i$ is constant. Substituting $A_i = \mathcal{V}(\mathfrak{s})/\Delta x_i$ we get

\begin{equation}
\text{div}(T) = \sum_{i=1}^{3} \lim_{\Delta x_i \to 0} \frac{1}{A_i} \int_{\Omega_i} \lim_{\Delta x_i \to 0} \left( \frac{T(x + \Delta x_i e_i) - T(x)}{\Delta x_i} \right) dA_i e_i
\end{equation}

Taking the limit of the average of the limit, as before, we arrive at the expression for the divergence in Cartesian coordinates:

\begin{equation}
\text{div}(T) = \frac{\partial T}{\partial x_i} e_i = \frac{\partial}{\partial x_i}(T e_i) = \frac{\partial T_i(x)}{\partial x_j} e_i
\end{equation}

It should be evident that all of the forms of the divergence of a tensor field given in Eqn. (90) are equivalent. The convenience of one form over another depends upon the application.

### Integral Theorems

**The divergence theorem.** There is an integration theorem worth mentioning here because it comes up repeatedly in solid mechanics. We call it the **divergence theorem** because it involves the divergence of a vector field. Consider again a region $\mathfrak{s}$ of arbitrary size and shape, with boundary $\Omega$ described by its normal vectors $n$. The divergence theorem can be stated as follows

\begin{equation}
\int_{\mathfrak{s}} \text{div} \mathbf{v} \, dV = \int_{\Omega} \mathbf{v} \cdot n \, dA
\end{equation}

This remarkable theorem, also known as Green's theorem or Gauss's theorem, relates an integral over the volume of a region to an integral over the boundary of that same region. It applies to any sufficiently well-behaved vector field $\mathbf{v}(x)$, and, thus, is very powerful. The proof of the divergence theorem can be carried out along many lines. The one in Schey (1973) is particularly descriptive. Schey's argument goes something as follows.

Partition the region $\mathfrak{s}$ into $N$ small subregions $\mathfrak{s}_i$ each having volume $\mathcal{V}(\mathfrak{s}_i)$, surfaces $\Omega_i$, and unit outward normal vector field $n_i$, as illustrated in Fig. 24. The surface of a certain subregion is the union of interior surfaces shared with adjacent subregions and (possibly) part of the original exterior surface $\Omega$. The normal vectors along a shared surface between two adjacent subregions point in opposite directions, as shown in the figure. Consequently, if we sum the fluxes over all of the subregions we get

\begin{equation}
\int_{\mathfrak{s}} \text{div} \mathbf{v} \, dV = \int_{\Omega} \mathbf{v} \cdot n \, dA
\end{equation}
In other words, the contributions of fluxes across the interior surfaces cancel each other out because there is only one $v$ at a given point on the surface (provided that $v$ is a continuous field) while one normal is the negative of the other.

Let us define the "almost divergence" of the vector field to be the finite ratio of flux to volume of subregion $\mathcal{B}_i$

$$\mathcal{D}_i[v] = \frac{1}{\psi(\mathcal{B}_i)} \int_{\Omega_i} v \cdot n_i \, dA_i$$

and observe that $\mathcal{D}_i[v] \to \text{div}(v)$ in the limit as $\psi(\mathcal{B}_i) \to 0$. Multiplying Eqn. (93) through by $\psi(\mathcal{B}_i)$ and summing over all $N$ subregions, we can see from Eqn. (92) that

$$\int_{\Omega} v \cdot n \, dA = \sum_{i=1}^{N} \mathcal{D}_i[v] \psi(\mathcal{B}_i)$$

This equation holds no matter how many subregions there are in the partition. As the number of partitions is taken larger and larger the size of the subregions shrinks. In the limit as $N \to \infty$ the discrete elements pass to their infinitesimal limits, that is, $\mathcal{D}_i[v] \to \text{div}(v)$ and $\psi(\mathcal{B}_i) \to dV$. The limit of the sum is the integral over the volume

$$\lim_{N \to \infty} \sum_{i=1}^{N} \mathcal{D}_i[v] \psi(\mathcal{B}_i) = \int_{\Omega} \text{div} v \, dV$$

thereby completing the proof.

The utility of defining the divergence with the intrinsic formula, Eqn. (78), should be evident from the proof of the divergence theorem. This proof might

**Figure 24** A region in three-dimensional space partitioned into subregions $\mathcal{B}_i$, each with volume $\psi(\mathcal{B}_i)$, surface $\Omega_i$, and unit outward normal vector field $n_i$. 
not have the level of rigor that a mathematician would like (the limiting process
and crossover to infinitesimals being the sloppiest point), but the geometric ba-
sis lends it a clarity that is more than adequate for our purposes here.

The divergence theorem holds for any vector field \( \mathbf{v}(x) \) that is well behaved.
A simple way to think about "well-behavedness" is to consider some of the bad
things that might happen on the way to the limit. In particular, any of the ob-
jects, like \( \partial_i [v] \), must exist for all possible subdivisions. If the vector field has
a singular point (\( v \to \infty \)), then eventually the subdivision process will en-
counter it, and, for the subdomains on whose boundaries the singularity lies,
\( \partial_i [v] \) is not defined. Similarly, if the field has a bounded jump along some sur-
face (where \( v^- \neq v^+ \) on opposites sides of the jump), then for those subdo-
mains that have a boundary on the jump surface, the fluxes will not cancel out.
Many of these pathologies can be treated by enhancing the integral theorems
with features that account for them. We do not have to worry about the patholo-
gies if our vector field \( \mathbf{v} \) and its divergence are continuous over the domain \( \mathcal{B} \)
and on the surface \( \Omega \).

**Example 8.** The divergence theorem for the gradient of a scalar field is

\[
\int_\Omega \nabla g \, dV = \int_\Omega g \, n \, dA
\]

where \( \mathcal{B} \) is a region with surface \( \Omega \) having unit outward normal vector field \( \mathbf{n} \).
Verify the relationship by applying it to the function \( g = x_1^2 + x_2^2 + x_3^2 \) defined
on a cylinder of unit radius and unit height, centered at the origin.

![Circular cylinder definition for Example 8.](image)

The integral of the gradient over the volume is best done in cylindrical coor-
dinates. Let \( x_1 = r \cos \theta \), \( x_2 = r \sin \theta \), and \( x_3 = z \). The gradient of \( g \) can be
computed as \( \nabla g = 2x = 2(r \mathbf{e}_\theta + z \mathbf{e}_z) \), where \( \mathbf{e}_\theta(\theta) = \cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2 \).
The volume integral can be carried out as follows:

\[
\int_\mathcal{B} \nabla g \, dV = \int_0^h \int_0^{2\pi} \int_0^R 2(r \mathbf{e}_\theta + z \mathbf{e}_z) \, r \, dr \, d\theta \, dz = \pi R^2 h^2 \mathbf{e}_3
\]

(Observe that the integral of \( \mathbf{e}_\theta(\theta) \) with respect to \( \theta \) from 0 to \( 2\pi \) is zero). The
surface has the following characteristics:
Bottom Surface: \( z = 0 \quad n = -e_3 \quad g = r^2 \)
Top Surface: \( z = h \quad n = +e_3 \quad g = r^2 + h^2 \)
Lateral Surface: \( r = R \quad n = e_r \quad g = R^2 + z^2 \)

The surface integral can be carried out as follows, noting that the integrand for the top and bottom surfaces reduces to \((r^2 + h^2)e_3 - r^2e_3 = h^2e_3\),

\[ \iint_{\Omega} g n \, dA = \int_0^{2\pi} \int_0^h h^2 e_3 \, r \, dr \, d\theta + \int_0^{2\pi} \left( R^2 + z^2 \right) e_r(\theta) R \, d\theta \, dz \]

\[ = \pi R^2 h^2 e_3 \]

Clearly, the volume and surface integrals have the same value, as the divergence theorem promises.

There are integral theorems for the gradient of a scalar field, the gradient of a vector field and a tensor field (see next section) that are analogous to the divergence theorem. The statements and proofs of these theorem are left as an exercise (Problem 46).

**Divergence theorem for tensor fields.** Any tensor field satisfies the following integral theorem (divergence theorem)

\[ \iint_{\Omega} \text{div} T \, dV = \iint_{\Omega} T n \, dA \quad (96) \]

where, as before, \( \mathcal{V}(\mathcal{B}) \) is the volume of the region \( \mathcal{B} \), \( \Omega \) is the surface of the region, and \( n \) is the unit normal vector field to the surface. Proof of the divergence theorem for tensor fields is left as an exercise (Problem 46).

**Additional Reading**


Problems

1. Compute the values of the following expressions
   (a) $\delta_{ij}$
   (b) $\delta_{ij} \delta_{ij}$
   (c) $C_{ij} \delta_{ik}$
   (d) $\delta_{ab} \delta_{bc} \delta_{cd} ... \delta_{xy} \delta_{yz}$ (enough terms to exhaust the whole alphabet)

2. Let two vectors, $u$ and $v$, have components relative to some basis as $u = (5, -2, 1)$ and $v = (1, 1, 1)$. Compute the lengths of the vectors and the angle between them. Find the area of the parallelogram defined by $u$ and $v$.

3. The vertices of a triangle are given by the position vectors $a$, $b$, and $c$. The components of these vectors in a particular basis are $a = (0, 0, 0)$, $b = (1, 4, 3)$, and $c = (2, 3, 1)$. Using a vector approach, compute the area of the triangle. Find the area of the triangle projected onto the plane with normal $n = (0, 0, 1)$. Find the unit normal vector to the triangle.

4. Let the coordinates of four points $a$, $b$, $c$ and $d$ be given by the following position vectors $a = (1, 1, 1)$, $b = (2, 1, 1)$, $c = (1, 2, 2)$, and $d = (1, 1, 3)$ in the coordinate system shown. Find vectors normal to planes $abc$ and $bcd$. Find the angle between those vectors. Find the area of the triangle $abc$. Find the volume of the tetrahedron $abcd$.

5. Demonstrate that $(u \times v) \cdot w = u_i v_j w_k \epsilon_{ijk}$ from basic operations on the base vectors.

6. Show that the triple scalar product is skew-symmetric with respect to changing the order in which the vectors appear in the product. For example, show that

$$(u \times v) \cdot w = -(v \times u) \cdot w$$

To generalize this notion, any cyclic permutation (e.g., $u, v, w \rightarrow w, u, v$) of the order of the vectors leaves the algebraic sign of the product unchanged, while any acyclic permutation (e.g., $u, v, w \rightarrow v, u, w$) of the order of the vectors changes the sign. How does this observation relate to swapping rows of a matrix in the computation of the determinant of that matrix?

7. Use the observation that $\|u - v\|^2 = (u - v) \cdot (u - v)$ along with the distributive law for the dot product to show that

$$u \cdot v = \frac{1}{2}(\|u\|^2 + \|v\|^2 - \|v - u\|^2)$$

8. Prove the Schwarz inequality, $|u \cdot v| \leq \|u\| \|v\|$. Try to prove this inequality without using the formula $u \cdot v = \|u\| \|v\| \cos \theta(u, v)$.

9. Show that $[u \otimes v]^T = v \otimes u$ using the definition of the transpose of a tensor and by demonstrating that the two tensors give the same result when acting on arbitrary vectors $a$ and $b$.

10. Show that the components of a tensor $T$ and its transpose $T^T$ satisfy $[T^T]_{ij} = [T]_{ji}$.
11. Show that the tensor $T^T T$ is symmetric.

12. Consider any two tensors $S$ and $T$. Prove the following:
   (a) $\det(T^T) = \det(T)$
   (b) $\det(ST) = \det(S) \det(T)$
   (c) $[ST]^T = T^T S^T$
   (d) $[ST]^{-1} = T^{-1} S^{-1}$

13. Consider two Cartesian coordinate systems, one with basis $\{e_1, e_2, e_3\}$ and the other with basis $\{g_1, g_2, g_3\}$. Let $Q_{ij} = e_i \cdot e_j$ be the cosine of the angle between $e_i$ and $e_j$.
   (a) Show that $g_i = Q_{ij} e_j$ and $e_i = Q_{ij} g_j$ relate the two sets of base vectors.
   (b) We can define a rotation tensor $Q$ such that $e_i = Q g_i$. Show that this tensor can be expressed as $Q = Q_{ij} [g_i \otimes g_j]$, that is, $Q_{ij}$ are the components of $Q$ with respect to the basis $[g_i \otimes g_j]$. Show that the tensor can also be expressed in the form $Q = [e_i \otimes g_i]$.
   (c) We can define a rotation tensor $Q^T$, such that $g_i = Q^T e_i$ (the reverse rotation from part (b)). Show that this tensor can be expressed as $Q^T = Q_{ij} [e_j \otimes e_i]$, that is, $Q_{ij}$ are the components of $Q^T$ with respect to the basis $[e_j \otimes e_i]$. Show that the tensor can also be expressed in the form $Q^T = [g_i \otimes e_i]$.
   (d) Show that $Q^T Q = I$, which implies that the tensor $Q$ is orthogonal.

14. The components of tensors $T$ and $S$ and the components of vectors $u$ and $v$ are

$$T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad S = \begin{bmatrix} 0 & -2 & 1 \\ 2 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Compute the components of the vector $Su$. Find the cosine of the angle between $u$ and $Su$. Compute the determinants of $T$, $S$, and $TS$. Compute $T^T T$ and $T^{-1} S^{-1}$.

15. Verify that, for the particular case given here, the components of the tensor $T$ and the components of its inverse tensor $T^{-1}$ are

$$T = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad T^{-1} \approx \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

16. Consider two bases: $\{e_1, e_2, e_3\}$ and $\{g_1, g_2, g_3\}$. The basis $\{g_1, g_2, g_3\}$ is given in terms of the base vectors $\{e_1, e_2, e_3\}$ as

$$g_1 = \frac{1}{\sqrt{5}}(e_1 + e_2 + e_3), \quad g_2 = \frac{1}{\sqrt{6}}(2e_1 - e_2 - e_3), \quad g_3 = \frac{1}{\sqrt{2}}(e_2 - e_3)$$

The components of the tensor $T$ and vector $v$, relative to the basis $\{e_1, e_2, e_3\}$ are

$$T = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Compute the components of the vector $Tv$ in both bases. Compute the nine values of $T_{ij} T_{jk} T_{kl}$ (i.e., for $i, l = 1, 2, 3$). Find the components of the tensor $[T + T^T]$. Compute $T_{ii}$. 
17. Consider two bases: \( \{e_1, e_2, e_3\} \) and \( \{g_1, g_2, g_3\} \), where
\[
\begin{align*}
g_1 &= e_1 + e_2 + e_3, \\
g_2 &= e_2 + e_3, \\
g_3 &= e_2 - e_3
\end{align*}
\]
Compute \( Q_{ij} \) for the given bases. Compute the value of \( Q_{ik}Q_{ij} \). Explain why the identity \( Q_{ik}Q_{ij} = \delta_{ij} \) does not hold in this case.

Now consider a vector \( v = e_1 + 2e_2 + 3e_3 \) and a tensor \( T \) given as
\[
T = [e_2 \otimes e_1 - e_1 \otimes e_2] + [e_3 \otimes e_1 - e_1 \otimes e_3] + [e_3 \otimes e_2 - e_2 \otimes e_3]
\]
Compute the components of the vector \( Tv \) in both bases, i.e., find \( v, \) and \( v, \) so that the following relationship holds \( Tv = v, e_i = v, g_i \). Find the cosine of the angle between the vector \( v \) and the vector \( Tv \). Find the length of the vector \( Tv \).

18. A general \( n \)-th order tensor invariant can be defined as follows
\[
f_n(T) \equiv T_{i_1i_2} T_{i_2i_3} \cdots T_{i_ni_1}
\]
where \( \{i_1, i_2, \ldots, i_n\} \) are the \( n \) indices. For example, when \( n = 2 \) we can use \( \{i, j\} \) to give \( f_2(T) = T_{ij} T_{ji} \) when \( n = 3 \) we can use \( \{i, j, k\} \) to give \( f_2(T) = T_{ij} T_{jk} T_{ki} \). Prove that \( f_n(T) \) is invariant with respect to coordinate transformation.

19. Use the Cayley-Hamilton theorem to prove that for \( n \geq 4 \) all of the invariants \( f_n(T) \), defined in Problem 18, can be computed from \( f_1(T) \), \( f_2(T) \), and \( f_3(T) \).

20. From any tensor \( T \) one can compute an associated deviator tensor \( T_{dev} \), which has the property that the deviator tensor has no trace, i.e., \( \text{tr}(T_{dev}) = 0 \). Such a tensor can be obtained from the original tensor \( T \) simply by subtracting \( \alpha = \frac{1}{3} \text{tr}(T) \) times the identity from the original tensor, i.e., \( T_{dev} = T - \alpha I \). Show that \( \text{tr}(T_{dev}) = 0 \). Show that the principal directions of \( T_{dev} \) and \( T \) are identical, but that the principal values of \( T_{dev} \) are reduced by an amount \( \alpha \) from those of the tensor \( T \).

21. Consider a tensor \( T \) that has all repeated eigenvalues \( \mu_1 = \mu_2 = \mu_3 = \mu \). Show that the tensor \( T \) must have the form \( T = \mu I \).

22. Prove that the product of a tensor with itself \( n \) times can be represented as
\[
T^n = \sum_{i=1}^{3} (\mu_i)^n n_i \otimes n_i
\]
Hint: Observe that \( [n_i \otimes n_i][n_j \otimes n_j] = \delta_{ij} [n_i \otimes n_j] \) (no summation implied).

23. Show that the determinant of the tensor \( T \) can be expressed as follows
\[
\text{det}(T) = \frac{1}{3} \text{tr}(T^3) - \frac{1}{2} I_T \text{tr}(T^2) + \frac{1}{6} (I_T)^3
\]
where \( I_T = \text{tr}(T) = T_{ii} \) is the first invariant of \( T \). Use the Cayley-Hamilton theorem.

24. A certain state of deformation at a point in a body is described by the tensor \( T \), having the components relative to a certain basis of
\[
\mathbf{T} = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 5 & 1 \\ 0 & 1 & 2 \end{bmatrix}
\]
Find the eigenvalues and eigenvectors of $T$. Show that the invariants of the tensor $T$ are the same in the given basis and in the basis defined by the eigenvectors for the present case.

25. Find the tensor $T$ that has eigenvalues $\mu_1 = 1$, $\mu_2 = 2$, and $\mu_3 = 3$ with two of the associated eigenvectors given by

$$n_1 = \frac{1}{\sqrt{2}} (e_1 + e_2), \quad n_2 = \frac{1}{3} (-2e_1 + 2e_2 + e_3)$$

Is the tensor unique (i.e., is there another one with these same eigenproperties)?

26. Find the tensor $T$ that has eigenvalues $\mu_1 = 1$, $\mu_2 = 3$, and $\mu_3 = 3$, with two of the associated eigenvectors given by

$$n_1 = \frac{1}{\sqrt{3}} (e_1 + e_2 + e_3), \quad n_2 = \frac{1}{\sqrt{2}} (-e_2 + e_3)$$

Are the eigenvectors unique?

27. A certain state of deformation at a point in a body is described by the tensor $T$, having the components relative to a certain basis of

$$T \sim 10^{-2} \begin{bmatrix} 14 & 2 & 14 \\ 2 & -1 & -16 \\ 14 & -16 & 5 \end{bmatrix}$$

Let the principal values and principal directions be designated as $\mu$ and $n$. Show that $n_1 = (-1, 2, 2)$ is a principal direction and find $\mu_1$. The second principal value is $\mu_2 = 9 \times 10^{-2}$, find $n_2$. Find $\mu_3$ and $n_3$ with as little computation as possible.

28. The equation for balance of angular momentum can be expressed in terms of a tensor $T$ and the base vectors $e_i$ as $e_i \times (Te_i) = 0$ (sum on repeated index implied). What specific conditions must the components of the tensor $T$ satisfy in order for this equation to be satisfied?

29. The tensor $R$ that operates on vectors and reflects them (as in a mirror) with unit normal $n$ is given by

$$R = I - 2n \otimes n$$

Compute the vector that results from $[RR]v$. Compute the length of the vector $Rv$ in terms of the length of $v$. What is the inverse of the tensor $R$? Compute the eigenvalues and eigenvectors of $R$.

30. Let $v(x)$ and $u(x)$ be two vector fields, and $T(x)$ be a tensor field. Compute the following expressions in terms of the components $(v_i, u_i, T_{ij})$ of these fields relative to the basis $\{e_1, e_2, e_3\}$: $\nabla(Tv)$, $\nabla(u \cdot Tv)$, $\nabla(Tv)$, and $u \otimes Tv$.

31. Evaluate the following expressions:

(a) $\nabla(\nabla(\nabla x \otimes x))$

(b) $\nabla(\nabla(\nabla x \div x))$

(c) $\nabla \left[ \| \nabla \| x \| \|^2 \right]$

(d) $\nabla(x \otimes \nabla(x \otimes x))$

(e) $\nabla(x \div x)$

(f) $\nabla(x \cdot \nabla(x \cdot x))$

where $x = x_1 e_1 + x_2 e_2 + x_3 e_3$ is the position vector in space and all derivatives are with respect to the coordinates $x_i$. 
32. Let \( \mathbf{v}(x) = (x_2 - x_3) \mathbf{e}_1 + (x_3 - x_1) \mathbf{e}_2 + (x_1 - x_2) \mathbf{e}_3 \). Evaluate the following expressions: \( \nabla \mathbf{v} \), \( \nabla (\mathbf{x} \cdot \mathbf{v}) \), \( \text{div} \left[ \mathbf{x} \otimes \mathbf{v} \right] \), and \( \nabla (\mathbf{x} \times \mathbf{v}) \), where \( \mathbf{x} = x_i \mathbf{e}_i \) is the position vector. Evaluate the expressions at the point \( \mathbf{x} = \mathbf{e}_1 + 2 \mathbf{e}_2 + \mathbf{e}_3 \).

33. Let \( \mathbf{v}(x) \) be given by the following explicit function

\[
\mathbf{v}(x) = (x_1^2 + x_2 x_3) \mathbf{e}_1 + (x_2^2 + x_1 x_3) \mathbf{e}_2 + (x_3^2 + x_1 x_2) \mathbf{e}_3
\]

where \( \mathbf{x} \) is the position vector of any point and has components \( \{x_1, x_2, x_3\} \) relative to the Cartesian coordinate system shown. The vector field is defined on the spherical region \( \mathcal{B} \) of unit radius as shown in the sketch. Give an explicit expression for the unit normal vector field \( \mathbf{n}(x) \) to the surface of the sphere. Compute the gradient of the vector field \( \mathbf{v}(x) \). Compute the product \( [\nabla \mathbf{v}] \mathbf{n} \), i.e., the gradient of the vector field acting on the normal vector. Compute the divergence of the vector field \( \mathbf{v}(x) \). Compute the integral of \( \text{div} \mathbf{v} \) over the volume of the sphere. Compute the integral of \( \mathbf{v} \cdot \mathbf{n} \) over the surface of the sphere.

34. Let \( \mathbf{v}(x) \) be a vector field given by the following explicit function

\[
\mathbf{v}(x) = (x_1 x_1 + x_2 x_2) \ln(x_1^2 + x_2^2)
\]

where \( \ln(\cdot) \) indicates the natural logarithm of \( (\cdot) \). The vector field is defined on the cylindrical region \( \mathcal{B} \) of height \( h \) and radius \( R \) as shown in the sketch. Give an expression for the unit normal vector field \( \mathbf{n}(x) \) to the cylinder (including the ends). Compute the divergence of the vector field \( \mathbf{v}(x) \) and the integral of \( \text{div} \mathbf{v} \) over the volume of the cylinder.

35. Consider the scalar field \( g(x) = (\mathbf{x} \cdot \mathbf{x})^2 \). Compute \( \text{div} \left[ \nabla \left( \text{div} \left[ \nabla g(x) \right] \right) \right] \).

36. Let \( \mathbf{v}(x) \) be given by the following explicit function

\[
\mathbf{v}(x) = (x_2 + x_3) \mathbf{e}_1 + (x_1 + x_3) \mathbf{e}_2 + (x_1 + x_2) \mathbf{e}_3
\]

where \( \mathbf{x} \) is the position vector of any point and has components \( \{x_1, x_2, x_3\} \) relative to the Cartesian coordinate system as shown. The vector field is defined on the ellipsoidal region \( \mathcal{B} \) whose surface is described by the equation \( g(\mathbf{x}) = 2x_1^2 + 2x_2^2 + 2x_3^2 - 4 = 0 \). Give an expression for the unit normal vector field \( \mathbf{n}(x) \) to the ellipsoid. Compute the gradient of the vector field \( \mathbf{v}(x) \). Compute the product \( [\nabla \mathbf{v}] \mathbf{n} \), i.e., the gradient of the vector field acting on the normal vector. Compute the divergence of the vector field \( \mathbf{v}(x) \).

37. Evaluate the expression \( \text{div} \left[ \nabla (\mathbf{x} \cdot \mathbf{A} \mathbf{x}) \right] \), where \( \mathbf{A} \) is a constant tensor (i.e., it does not depend upon \( \mathbf{x} \)), and the vector \( \mathbf{x} \) has components \( \mathbf{x} = x_i \mathbf{e}_i \). The derivatives are to be taken with respect to the independent variables \( x_i \). Express the results in terms of the components of \( \mathbf{A} \) and \( \mathbf{x} \).

38. Let \( g(x) = e^{-\|\mathbf{x}\|^2} \) be a scalar field in three-dimensional space, where \( \| \mathbf{x} \| \) is the distance from the origin to the point \( \mathbf{x} \). Qualitatively describe the behavior of the function (a one- or two-dimensional analogy might be helpful). Compute the gradient \( \nabla g \) of the field. Where does the gradient of the function go to zero?
39. Consider a tensor field $T$ defined on a tetrahedral region bounded by the coordinate planes $x_1 = 0, x_2 = 0, x_3 = 0$, and the oblique plane $6x_1 + 3x_2 + 2x_3 = 6$, as shown in the sketch. The tensor field has the particular expression $T = b \otimes x$, where $b$ is a constant vector and $x$ is the position vector $x = x_i e_i$. Compute the integral of $\text{div}(T)$ over the volume and the integral of $Tn$ over the surface of the tetrahedron (and thereby show that they give the same result, as promised by the divergence theorem). Note that the volume of the tetrahedron of the given dimensions is one.

40. Let $v(x) = x$ on a spherical region of radius $R$, centered at the origin. Compute the integral of $\text{div}(v)$ over the volume of the sphere and compute the integral of the flux $v \cdot n$, where $n$ is the unit normal to the sphere, over the surface of the sphere. Give the result in terms of the radius $R$. What does this calculation tell you about the ratio of surface area to volume of a sphere?

41. The Laplacian of a scalar field is a scalar measure of the second derivative of the field, defined as $\nabla^2 g(x) \equiv \nabla \cdot (\nabla g(x))$. Write the component (index) form of the Laplacian of $g$ in Cartesian coordinates. Compute the Laplacian of the scalar field of Problem 38.

42. Compute $\text{div}(T)$, where $T(x) = (x \cdot x)I - 2x \otimes x$ is a tensor field.

43. Let $u(x), v(x)$, and $w(x)$ be vector fields and let $T(x)$ be a tensor field. Compute the component forms of the following derivatives of products of vectors

(a) $\nabla (u \cdot v)$
(b) $\nabla (u \times v)$
(c) $\nabla u \cdot v$
(d) $\text{div}(Tv)$
(e) $\nabla (u \cdot Tv)$
(f) $\nabla (Tv)$
(g) $\text{div}(u \otimes v)$
(h) $\text{div}([u \otimes v]w)$
(i) $\nabla ([u \times v] \cdot w)$

44. Use the same reasoning that was used to derive the three-dimensional version of the divergence theorem to develop (a) a one-dimensional version, and (b) a two-dimensional version of the theorem. Use sketches to illustrate your definitions and draw any possible analogies with the three-dimensional case.

45. Consider a vector field $v(x)$ on a region $\Omega$ with surface $\Omega$ having unit normal field $n$. The “curl” of the vector field can be defined as

$$\text{curl}(v) \equiv \lim_{\Omega(\mathcal{V}) \to 0} \frac{1}{\mathcal{V}(\mathcal{B})} \int_{\Omega} v \times n \, dA$$

Show (using the cuboid for $\mathcal{B}$, as in the text) that the expression for $\text{curl}(v)$ is

$$\text{curl}(v) = \frac{\partial v}{\partial x_i} e_i = \left( \frac{\partial v_2}{\partial x_3} - \frac{\partial v_3}{\partial x_2} \right) e_1 + \left( \frac{\partial v_3}{\partial x_1} - \frac{\partial v_1}{\partial x_3} \right) e_2 + \left( \frac{\partial v_1}{\partial x_2} - \frac{\partial v_2}{\partial x_1} \right) e_3$$

Note that many authors define the curl to be the negative of the definition given here, which is easily achieved by using the flux $n \times v$ instead. The form presented here seems to be more consistent with our other definitions of derivatives of vector fields.

46. Consider variously a scalar field $g(x)$, a vector field $v(x)$, and a tensor field $T(x)$ on a region $\mathcal{B}$ with surface $\Omega$ with unit normal vector field $n$. Prove the following theorems
\[ \int_{\Omega} \nabla g \, dV = \int_{\Omega} g \, \mathbf{n} \, dA, \quad \int_{\Omega} \nabla v \, dV = \int_{\Omega} v \otimes \mathbf{n} \, dA, \quad \int_{\Omega} \text{div} \mathbf{T} \, dV = \int_{\Omega} \mathbf{T} \, \mathbf{n} \, dA \]

47. Use the divergence theorem for a vector field to show the following identities

(a) \textit{Green's first identity} for scalar functions \( u(x) \) and \( v(x) \), (Hint: Let \( v(x) = u \nabla v \))

\[ \int_{\Omega} \left( u \nabla^2 v + \nabla u \cdot \nabla v \right) \, dV = \int_{\Omega} \mathbf{n} \cdot (u \nabla v) \, dA \]

(b) \textit{Green's second identity} for scalar functions \( u(x) \) and \( v(x) \), (Hint: Let \( v(x) = u \nabla v - v \nabla u \))

\[ \int_{\Omega} \left( u \nabla^2 v - v \nabla^2 u \right) \, dV = \int_{\Omega} \mathbf{n} \cdot (u \nabla v - v \nabla u) \, dA \]

48. Many problems are more conveniently formulated and solved in cylindrical coordinates \((r, \theta, z)\). In cylindrical coordinates, the components of a vector \( \mathbf{v} \) can be expressed as

\[ \mathbf{v}(r, \theta, z) = v_r \mathbf{e}_r + v_\theta \mathbf{e}_\theta + v_z \mathbf{e}_z \]

where the components \( v_r, v_\theta, \) and \( v_z \) are each functions of the coordinates \((r, \theta, z)\). However, now the base vectors \( \mathbf{e}_r(\theta) \) and \( \mathbf{e}_\theta(\theta) \) depend upon the coordinate \( \theta \). We must account for this dependence of the base vectors on the coordinates when computing derivatives of the vector field.

Using the coordinate-free definition of the divergence of a vector field, Eqn. (78), show that the divergence of \( \mathbf{v} \) in cylindrical coordinates is given by

\[ \text{div} \mathbf{v}(r, \theta, z) = \frac{1}{r} \frac{\partial}{\partial r} (r v_r) + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z} \]

(Hint: Observe from the figure that \( \mathbf{n}_1 = \mathbf{e}_\theta(\theta + \Delta \theta) \) and \( \mathbf{n}_2 = -\mathbf{e}_\theta(\theta) \) and are constant over the faces 1 and 2, respectively. The normal vectors \( \mathbf{n}_3 = \mathbf{e}_r(\xi) \) and \( \mathbf{n}_4 = -\mathbf{e}_r(\xi) \), with \( \xi \in [\theta, \theta + \Delta \theta] \), vary over faces 3 and 4. Finally, note that \( \mathbf{n}_5 = \mathbf{e}_z \) and \( \mathbf{n}_6 = -\mathbf{e}_z \) are constant over faces 5 and 6.)

Note that the volume of the wedge is \( V(\mathcal{B}) = r \Delta \theta \Delta r \Delta z \) plus terms of higher order that vanish more quickly in the limit as \( V(\mathcal{B}) \to 0 \).