Notes for Introduction to Continuum Mechanics

Lynn Schreyer Bennethum

February 1, 2006

Contents

Chapter 1

Vector Calculus, Tensors, and Indicial Notation

In this chapter we review vectors (which students should have seen in Calculus 3 and/or undergraduate physics), introduce tensors (a generalization of vectors to higher order), and introduce indicial notation, a convenient notation for manipulating tensors, vectors, and equations involving these quantities. Indicial notation, sometimes referred to as Einstein notation, was invented by (who else?) Einstein. It is a way of writing vector equations using indices in such a way that it represents clearly which parts of a product of (first-order, second-order, third-order) tensors are being acted on/upon and added together. It also give a quick visual cue as to whether an equation makes sense.

We proceed with a disclaimer: For all but the last section of this chapter the indicial notation presented is valid only for an orthonormal coordinate system. For cylindrical or spherical coordinates we need a more general form of the indicial notation and this is introduced in the last section.

We begin with the definition of vector, and then introduce indicial notation for a vector.

Definition 1.1 *A* **vector** *is an object representing a direction and magnitude.*

Thus for example, a scalar, a, is not a vector since although it has a magnitude ($|a|$), it does not represent a direction.

Note that the definition of a vector is coordinate-system independent. We can denote the vector in *direct notation* as $\boldsymbol{v}$, where in 2-dimensional Cartesian coordinates the vector $\boldsymbol{v}$ has components (v_1, v_2) and in 3-dimensions it has components (v_1, v_2, v_3). Note that the components are coordinate dependent: the components change depending upon the coordinate system being used. Direct notation, $\boldsymbol{v}$, is coordinate independent. In *indicial notation*, the components of the vector are denoted as v_i where i is the index, and it is assumed that $i = 1, 2, 3$ unless otherwise noted. We will make this rule number one:

> **Indicial Notation Rule Number 1:** *Unless otherwise noted, all indices take on the values 1, 2, and 3.*

Thus $\boldsymbol{w}$ can be represented as w_i or w_j. Note that the free index acts just as a variable in a function. Whereas $f(x) = x^2$ is the same function as $f(y) = y^2$, w_i represents the same vector as w_j.

1.1 Dot Product

The dot product of two vectors, $\boldsymbol{u}$ and $\boldsymbol{v}$, in direct notation is given by $\boldsymbol{u} \cdot \boldsymbol{v}$. Because the dot product is related to the angle between the two vectors, θ,

$$\boldsymbol{u} \cdot \boldsymbol{v} = ||\boldsymbol{u}||\ ||\boldsymbol{v}|| \cos\theta \tag{1.1}$$

it is a quantity which does not vary with the coordinate system. In indicial notation the dot product is:

$$u_i v_i = \sum_{i=1}^{3} u_i v_i = u_1 v_1 + u_2 v_2 + u_3 v_3. \tag{1.2}$$

In indicial notation the summation on i is implied. Thus:

Indicial Notation Rule Number 2: *Repeated indices implies summation.*

We note that with repeated indices, the index is a dummy variable so that $u_i v_i = u_j v_j$ in much the same way that $\int f(x)\, dx = \int f(y)\, dy$.

Next we review some definitions.

Definition 1.2 *If* $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ *then the vectors are* **orthogonal**.

Definition 1.3 *If* $\boldsymbol{u} \cdot \boldsymbol{u} = 1$ *then* $\boldsymbol{u}$ *is a* **unit vector**.

Definition 1.4 *If* $\boldsymbol{u} \cdot \boldsymbol{u} = 1$, $\boldsymbol{v} \cdot \boldsymbol{v} = 1$, *and* $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ *then the set* $\{\boldsymbol{u}, \boldsymbol{v}\}$ *is an* **orthonormal set**.

Exercise: What conditions are necessary for the set $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ to be an orthonormal set?

The dot product is often used in the *decomposition* of a vector $\boldsymbol{v}$. Given a unit vector $\boldsymbol{n}$, a vector $\boldsymbol{v}$ can be decomposed in such a way that

$$\boldsymbol{v} = \boldsymbol{v}^p + \boldsymbol{v}^o$$

where $\boldsymbol{v}^p$ is parallel to $\boldsymbol{n}$ and $\boldsymbol{v}^o$ is orthogonal to $\boldsymbol{n}$. The vector $\boldsymbol{v}^p$ is called the *projection* of $\boldsymbol{v}$ onto $\boldsymbol{n}$ and is given by:

$$\boldsymbol{v}^p = (\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n} \qquad v_j^p = v_i n_i n_j. \tag{1.3}$$

In the indicial form of the above equation j is the free index in each term because it appears exactly once in each term. And this leads us to the next indicial notation rule:

Indicial Notation Rule Number 3: *In each term of an equation, the free indices must match.*

So for example, $u_i + v_i = w_i$ is correct, but $u_i + v_j = w_i$ is not.

Often one wants to project a vector $\boldsymbol{v}$ onto a not necessarily unit vector $\boldsymbol{w}$. This can be done if one replaces $\boldsymbol{n}$ in the above formula by $\dfrac{\boldsymbol{w}}{||\boldsymbol{w}||}$, which is a unit vector in the direction of $\boldsymbol{w}$. The resulting formula is what is often given in the third semester of calculus: $\boldsymbol{v}^p = \dfrac{\boldsymbol{v} \cdot \boldsymbol{w}}{|\boldsymbol{w}|^2}\,\boldsymbol{w}$.

The component vector of $\boldsymbol{v}$ perpendicular to $\boldsymbol{n}$ is

$$\boldsymbol{v}^o = \boldsymbol{v} - \boldsymbol{v}^p = \boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n} = \boldsymbol{v} \cdot (\boldsymbol{I} - \boldsymbol{n}\boldsymbol{n}) \tag{1.4}$$

$$v_j = v_j^p - v_i n_i n_j. \tag{1.5}$$

Note that we don't need parenthesis in indicial notation. This is one of the properties of indicial notation: *Order doesn't matter*, since the order of multiplication is determined by the indices.

1.2 Tensors

We need the definition of a vector space to begin:

Definition 1.5 *A* **vector space***, V, consists of a set of vectors for which vector addition and scalar multiplication are defined and which satisfy the following properties for all vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in V:*

1. *If $\boldsymbol{u}$ and $\boldsymbol{v}$ are any two vectors in V, then $\boldsymbol{u} + \boldsymbol{v}$ is also in V (closure under additivity).*

2. *If k is any real number and $\boldsymbol{u}$ is any vector in V then $k\boldsymbol{u}$ is also in V (closure under scalar multiplication).*

3. *$\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (commutativity).*

4. *$\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ (associativity of vector addition).*

5. *There is a vector, $\boldsymbol{0}$ such that for any $\boldsymbol{u}$ in V, $\boldsymbol{0} + \boldsymbol{u} = \boldsymbol{u}$ (additive identity).*

6. *For any $\boldsymbol{u}$ in V there is another vector, $-\boldsymbol{u}$ such that $\boldsymbol{u} + (-\boldsymbol{u}) = (-\boldsymbol{u}) + \boldsymbol{u} = \boldsymbol{0}$ (existence of an additive inverse).*

7. *$k(l\boldsymbol{u}) = (kl)(\boldsymbol{u})$ (associativity of scalr multipliation).*

8. *$(k + l)\boldsymbol{u} = k\boldsymbol{u} + l\boldsymbol{u}$ (distributivity for scalar sums).*

9. *$k(\boldsymbol{u} + \boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$ (distributivity for vector sums).*

10. *$1\boldsymbol{u} = \boldsymbol{u}$ (scalar multiplicative identity).*

A tensor is a generalization of a vector in the following sense. We consider a vector as a linear transformation acting on a vector space: through the dot product a vector is a linear transformation which maps a vector to a scalar: $\boldsymbol{v} \cdot (a\boldsymbol{u} + \boldsymbol{w}) = a\boldsymbol{v} \cdot \boldsymbol{u} + \boldsymbol{v} \cdot \boldsymbol{w}$. Then a generalization gives us the definition of a second-order tensor:

Definition 1.6 *A* **second-order tensor** *is defined to be a linear tranformation from a vector space onto itself, i.e., if* $\boldsymbol{v}$ *and* $\boldsymbol{w}$ *are arbitrary vectors in the vector space* V *then,* $\underline{\underline{\boldsymbol{B}}}$ *is a second-order tensor if* $\underline{\underline{\boldsymbol{B}}}(c\boldsymbol{u} + \boldsymbol{v}) = c\underline{\underline{\boldsymbol{B}}}(\boldsymbol{u}) + \underline{\underline{\boldsymbol{B}}}(\boldsymbol{v})$.

As long as every vector in the vector space can be represented as a linear combination of the basis vectors, then it can be proven that all second-order tensors can be represented by a matrix. In particular, matrices are specific representations of second-order tensors and the representation is dependent on the coordinate system. From linear algebra it is known that rotations, reflections, and projections are examples of linear transformations and can all be represented by matrices for a particular coordinate system. These are examples of second-order tensors. Another example is the transformation $B(\boldsymbol{u}) = \boldsymbol{b} \times \boldsymbol{u}$ where $\boldsymbol{b}$ is a given vector. Though this is a linear transformation mapping a vector space onto itself (and so it is a second-order tensor), it is not readily apparent what its matrix representation is.

In linear algebra, a linear system of equations is usually written as $\boldsymbol{w} = \boldsymbol{B}\boldsymbol{u}$. It turns out that this notation will not be precise enough once higher-order tensors are defined. So in direct tensor notation, this system of equations is written as: $\boldsymbol{w} = \underline{\underline{\boldsymbol{B}}} \cdot \boldsymbol{u}$, which in indicial notation is

$$w_i = B_{ij}u_j = u_jB_{ij}. \tag{1.6}$$

Order doesn't matter in indicial notation since repeated indices indicate what is being summed. Note that the free indices match. In general a single dot indicates one index is repeated.

Exercise:

Matrix Notation	Direct Notation	Indicial Notation
$\boldsymbol{w} = \boldsymbol{B}^T\boldsymbol{u}$	$\boldsymbol{w} = \underline{\underline{\boldsymbol{B}}}^T \cdot \boldsymbol{u}$	$w_j = B_{ij}u_i$
$\boldsymbol{w}^T = \boldsymbol{u}^T\boldsymbol{B}$	$\boldsymbol{w} = \boldsymbol{u} \cdot \underline{\underline{\boldsymbol{B}}}$	(i)
$a = \boldsymbol{u}^T\boldsymbol{B}\boldsymbol{v}$	(ii)	$a = B_{ij}u_iv_j$

Ans:

$(i) w_j = u_iB_{ij} = B_{ij}u_i, \quad (ii) a = \boldsymbol{u} \cdot \underline{\underline{\boldsymbol{B}}} \cdot \boldsymbol{v}.$

As a side remark: direct notation can be quite different from matrix notation. As an example, the direct notation expression $\boldsymbol{w} = \boldsymbol{u} \cdot \underline{\underline{\boldsymbol{B}}}$ would be written as $\boldsymbol{w}^T = \boldsymbol{u}^T\boldsymbol{B}$ in matrix notation. In matrix notation no dot, $\cdot$, is used to denote multiplication of a vector and a matrix, as it is assumed that wherever a matrix and a vector appear next to each other, the two are multiplied. Further it is generally understood which quantities are matrices and what are vectors so that no distinguishing marks are used to denote a matrix. In tensor notation, the general convention is that vectors are denoted by lower-case bold letters, and matrices are denoted by capital bold-face letters. In this text, to further distinguish between second-order and higher-order tensors, two underlines will be used to denote a second-order tensor. In direct notation a transpose on vectors is sometimes used to denote that the components should be written horizontally rather than vertically.

We now formalize our notation:

Object	Direct	Indicial
scalar	a	a
vector	$\boldsymbol{u}$	u_i
2nd-order tensor	$\underline{\underline{\boldsymbol{B}}}$	B_{ij}
3rd-order tensor	$\underline{\underline{\underline{\boldsymbol{C}}}}$	C_{ijk}
4th-order tensor	$\underline{\underline{\underline{\underline{\boldsymbol{D}}}}}$	D_{ijkl}

A third-order tensor is also a linear operator. It can send vectors into second-order tensors vis $\underline{\underline{\underline{\boldsymbol{C}}}} \cdot \boldsymbol{v} = \underline{\underline{\boldsymbol{B}}}$ or $C_{ijk}v_k = B_{ij}$, or it can send second-order tensors to third-order tensors: $\underline{\underline{\underline{\boldsymbol{C}}}} \cdot \underline{\underline{\boldsymbol{B}}} = \underline{\underline{\underline{\boldsymbol{D}}}}$ or $C_{ijk}B_{kl} = D_{ijl}$. This last operation, $\underline{\underline{\underline{\boldsymbol{C}}}} \cdot \underline{\underline{\boldsymbol{B}}}$, can be interpreted as a third order tensor acting on a series of vectors, and then the resulting second-order tensor acts on a vector. Since by definition second-order tensors are linear transformations, this operation is also linear. We can even define a *double contraction*

$$\underline{\underline{\underline{\boldsymbol{C}}}} : \underline{\underline{\boldsymbol{B}}} = \boldsymbol{v} \qquad\qquad C_{ijk}B_{jk} = v_i, \tag{1.7}$$

so that a third-order tensor is also a linear transformation mapping a second-order tensor to a vector.

In continuum mechanics, it is very important to distinguish between scalars, vectors, and higher order tensors. Note that $0 \neq \mathbf{0} \neq \underline{\underline{\mathbf{0}}} \neq ....$

Exercise:

$$u_i + v_i = w_{\underline{(i)}} \tag{1.8}$$
$$u_i + v_j = w_{\underline{(ii)}} \tag{1.9}$$
$$A_{ij}b_i + u_{\underline{(iii)}} = w_{\underline{(iv)}} \tag{1.10}$$
$$C_{ijk}B_{kl}u_i + E_{jlmi}T_{im} = Q_{\underline{(v)}} \tag{1.11}$$

Ans:
(i) i, (ii) Not defined, (iii) j , (iv) j, (v) jl or lj. Note that the solution (v) has two possible answers. This will be addressed shortly.

The following rule is a good check to be sure equation manipulations involving tensors are valid:

Indicial Notation Rule Number 4: *Indices repeated more than two times in a single term are not allowed.*

So for example $C_{ijk}B_{kl}u_k$ is not allowed. This is due to the fact that there is no corresponding direct notation (which is coordinate independent), and thus violates the spirit of tensors. In addition, the appearance of such a term usually indicates an unphysical quantity (since physical quantities such as forces, heat fluxes, and strains and stresses represent the same thing regardless of the coordinate system). If this is indeed the sort of quantity that must be represented, then the summation sign should be used.

1.3 Special Matrices/Second-Order Tensors

Any second-order tensor (and hence matrix), $\underline{\underline{\boldsymbol{A}}}$, can be decomposed into a symmetric part and an anti-symmetric part.

Definition 1.7 *A second-order tensor,* $\underline{\underline{\boldsymbol{A}}}$*, is* **symmetric** *if* $\underline{\underline{\boldsymbol{A}}}^T = \underline{\underline{\boldsymbol{A}}}$.

Definition 1.8 *A second-order tensor,* $\underline{\underline{\boldsymbol{A}}}$*, is* **anti-symmetric** *if* $\underline{\underline{\boldsymbol{A}}}^T + \underline{\underline{\boldsymbol{A}}} = \underline{\underline{\boldsymbol{0}}}$.

Any second-order tensor $\underline{\underline{\boldsymbol{A}}}$ can be decomposed into a symmetric part, $\underline{\underline{\boldsymbol{A}}}^s$, and an anti-symmetric part, $\underline{\underline{\boldsymbol{A}}}^a$:

$$\underline{\underline{\boldsymbol{A}}}^s = \frac{1}{2}\left(\underline{\underline{\boldsymbol{A}}} + \underline{\underline{\boldsymbol{A}}}^T\right) \qquad \underline{\underline{\boldsymbol{A}}}^a = \frac{1}{2}\left(\underline{\underline{\boldsymbol{A}}} - \underline{\underline{\boldsymbol{A}}}^T\right) \tag{1.12}$$

so that

$$\underline{\underline{\boldsymbol{A}}} = \underline{\underline{\boldsymbol{A}}}^s + \underline{\underline{\boldsymbol{A}}}^a. \tag{1.13}$$

The **identity matrix**,

$$\underline{\underline{\boldsymbol{I}}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.14}$$

is, in indicial notation:

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}, \tag{1.15}$$

where δ_{ij} is the *Kronecker delta*, a special symbol used to represent the identity tensor. Using the summation rules of indicial notation and the definition of the Kronecker delta, we see that

$$\delta_{ij}b_j = \left\{ \begin{array}{l} \delta_{11}b_1 + \delta_{12}b_2 + \delta_{13}b_3 = b_1 \\ \delta_{21}b_1 + \delta_{22}b_2 + \delta_{23}b_3 = b_2 \\ \delta_{31}b_1 + \delta_{32}b_2 + \delta_{33}b_3 = b_3 \end{array} \right\} = b_i \tag{1.16}$$

which in direct notation just says that $\underline{\underline{\boldsymbol{I}}} \cdot \boldsymbol{b} = \boldsymbol{b}$. Whenever one is simplifying a term with the Kronecker delta, the idea is to replace one of the indices of δ_{ij} with the other and drop the Kronecker delta. Often it does not matter which index is replaced, but if it involves a free index, the free index must be kept.

Exercise:

1.
$$A_{ijk}\delta_{ik} = \underline{(i)}$$

2.
$$T_{ij}\delta_{ij} = \underline{(ii)}$$

3. If $\{\boldsymbol{e}^1,\ \boldsymbol{e}^2,\ \boldsymbol{e}^3\}$ is an orthonormal set, then $\boldsymbol{e}^i \cdot \boldsymbol{e}^j = \underline{(iii)}$.

Ans:
(i): $A_{iji} = A_{kjk}$ (ii): $T_{ij}\delta_{ij} = T_{ii} = \mathrm{tr}(\underline{\underline{\boldsymbol{T}}})$ (iii) δ_{ij}.

1.4 Contraction and Tensor Products

The dot product is an example of a contraction, i.e. a repeated index. This is called a *contraction* as it reduces the size of the tensor - e.g. a dot product takes 2 first-order tensors and produces a scalar. As a review:

Direct	Indicial
$\underline{\underline{\boldsymbol{A}}} \cdot \boldsymbol{v}$	$A_{ij}v_j$
$\underline{\underline{\boldsymbol{A}}} \cdot \underline{\underline{\boldsymbol{B}}}$	$A_{ij}B_{jk}$
$\underline{\underline{\boldsymbol{A}}} : \underline{\underline{\boldsymbol{B}}}$	$A_{ij}B_{ij}$

Some authors use the following notation: $\underline{\underline{\boldsymbol{A}}} \cdot\cdot \underline{\underline{\boldsymbol{B}}} = A_{ij}B_{ji}$ although this seems rare.

The **tensor product** does not reduce the order of the tensors. We define:

$$\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^T = \begin{bmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{bmatrix} \tag{1.17}$$

and in indicial notation $\boldsymbol{u} \otimes \boldsymbol{v}$ is represented as u_iv_j. So for example

$$\boldsymbol{e}^1 \otimes \boldsymbol{e}^2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{1.18}$$

and similarly

$$\boldsymbol{e}^2 \otimes \boldsymbol{e}^1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{1.19}$$

In particular we notice that in general $\boldsymbol{e}^i \otimes \boldsymbol{e}^j$ gives a matrix of all zero's except for a 1 in the i, j^{th} position, and the tensor product is not commutative.

In keeping with the spririt of the definition of a second-order tensor as a linear operator, one may formally define the tensor product as a linear operator acting on a vector space V. If $\boldsymbol{w}$ is any vector in V then

$$(\boldsymbol{u} \otimes \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u}(\boldsymbol{v} \cdot \boldsymbol{w}).$$

Thus in order to determine $(\boldsymbol{u} \otimes \boldsymbol{v}) \cdot (\boldsymbol{w} \otimes \boldsymbol{x})$ we look at its action on a vector $\boldsymbol{a}$:

$$\begin{aligned}
(\boldsymbol{u} \otimes \boldsymbol{v}) \cdot (\boldsymbol{w} \otimes \boldsymbol{x})\boldsymbol{a} &= \boldsymbol{u} \otimes \boldsymbol{v} \cdot \{(\boldsymbol{w} \otimes \boldsymbol{x})\boldsymbol{a}\} \\
&= (\boldsymbol{u} \otimes \boldsymbol{v}) \cdot \boldsymbol{w}(\boldsymbol{x} \cdot \boldsymbol{a}) \\
&= \boldsymbol{u}\boldsymbol{v} \cdot \boldsymbol{w}(\boldsymbol{x} \cdot \boldsymbol{a}) \\
&= \boldsymbol{v} \cdot \boldsymbol{w}(\boldsymbol{u} \otimes \boldsymbol{x}) \cdot \boldsymbol{a}
\end{aligned}$$

so that we have

$$(\boldsymbol{u} \otimes \boldsymbol{v}) \cdot (\boldsymbol{w} \otimes \boldsymbol{x}) = \boldsymbol{v} \cdot \boldsymbol{w}\boldsymbol{u} \otimes \boldsymbol{x},$$

or in indicial notation

$$u_i v_j w_j x_k = v_j w_j u_i x_k.$$

In general this formal definition of a tensor product is combersome, and so we will make use of this definition only if necessary.

Most of the time the operator $\otimes$ is omitted, so that the above expression may be written as

$$(\boldsymbol{u} \otimes \boldsymbol{v}) \cdot (\boldsymbol{w} \otimes \boldsymbol{x}) = \boldsymbol{u}\boldsymbol{v} \cdot \boldsymbol{w}\boldsymbol{x}.$$

Note that the tensor product, $\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}$, implies two, non-repeating, indices, so that no summation is implied. The result is a second-order tensor ($u_i v_j$, which has two free indices).

Now consider $\boldsymbol{e}^i \otimes \boldsymbol{e}^j$, which consists of 9 elements and which forms a basis for matrices in Cartesian coordinates. The matrix $\boldsymbol{e}^i \otimes \boldsymbol{e}^j$ consist of all zeros except in the (i, j) position, where it contains a 1. So just as we can write a vector as:

$$\boldsymbol{v} = v_i \boldsymbol{e}^i$$

we can write a second-order tensor as

$$\underline{\underline{\boldsymbol{A}}} = A_{ij} \boldsymbol{e}^i \otimes \boldsymbol{e}^j.$$

This allows us to distinguish between a second order tensor and its transpose if necessary:

$$\begin{aligned}
\underline{\underline{\boldsymbol{A}}} &= A_{ij} \boldsymbol{e}^i \otimes \boldsymbol{e}^j \\
\underline{\underline{\boldsymbol{A}}}^T &= A_{ji} \boldsymbol{e}^i \otimes \boldsymbol{e}^j.
\end{aligned}$$

Writing the basis vectors can be a bit cumbersome, and consequently the basis vectors are usually written only when necessary, such as to dinstinguish between a tensor and its transpose or when working in a curvi-linear coordinate system.

Exercise:
Write the following indicial notation expressions in direct notation.

1. $$A_{ij}B_{jk}\boldsymbol{e}^i \otimes \boldsymbol{e}^k = \underline{\underline{\boldsymbol{A}}} \cdot \underline{\underline{\boldsymbol{B}}}$$

2. $$A_{ij}B_{ik}\boldsymbol{e}^j \otimes \boldsymbol{e}^k = \underline{(i)}$$

3. $$A_{ij}B_{kj}\boldsymbol{e}^i \otimes \boldsymbol{e}^k = \underline{(ii)}$$

4. $$A_{ij}B_{ki}\boldsymbol{e}^j \otimes \boldsymbol{e}^k = \underline{(iii)}$$

5. $$A_{ij}B_{ki}\boldsymbol{e}^k \otimes \boldsymbol{e}^j = \underline{(iv)}$$

Ans:
(i): $\underline{\underline{\boldsymbol{A}}}^T \cdot \underline{\underline{\boldsymbol{B}}}$ (ii): $\underline{\underline{\boldsymbol{A}}} \cdot \underline{\underline{\boldsymbol{B}}}^T$ (iii): $\underline{\underline{\boldsymbol{A}}}^T \cdot \underline{\underline{\boldsymbol{B}}}^T$ (iv): $\underline{\underline{\boldsymbol{B}}} \cdot \underline{\underline{\boldsymbol{A}}}$

1.5 Cross Product and Determinant

Recall that one goal is to develop a system which is coordinate independent. The dot product is related to the angle between two vectors. A second way of "multiplying" two vectors in a physically (or geometrically) meaningful way is the cross product. One of the most important properties of the cross product is that it give a vector orthogonal to both vectors. To calculate the cross product, one takes the determinant of the matrix in which the first row consists of $\boldsymbol{i}$, $\mathbf{j}$, and $\boldsymbol{k}$, the second row consists of the components of the first vector, and the third row consists of the components of the second vector. For example if:

$$\boldsymbol{u} = (1, 2, 3) \qquad \boldsymbol{v} = (-1, 1, 0)$$

then

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \mathbf{j} & \boldsymbol{k} \\ 1 & 2 & 3 \\ -1 & 1 & 0 \end{vmatrix} = \boldsymbol{i} \begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix} - \mathbf{j} \begin{vmatrix} 1 & 3 \\ -1 & 0 \end{vmatrix} + \boldsymbol{k} \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = -3\boldsymbol{i} - 3\mathbf{j} + 3\boldsymbol{k}.$$

There are several properties of the cross product:

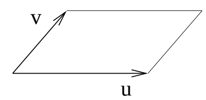
1. $\boldsymbol{u} \times \boldsymbol{v}$ gives a vector orthogonal to both $\boldsymbol{u}$ and $\boldsymbol{v}$.

2. The direction of $\boldsymbol{u} \times \boldsymbol{v}$ is determined by the right-hand rule (point your right-hand fingers in the direction of $\boldsymbol{u}$, curl them in the direction of $\boldsymbol{v}$, and the direction of your thumb is the direction of $\boldsymbol{u} \times \boldsymbol{v}$). This gives us the next property.

3. $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$.

4. $|\boldsymbol{u} \times \boldsymbol{v}| = |\boldsymbol{u}||\boldsymbol{v}| \sin\theta$ where θ is the acute angle between $\boldsymbol{u}$ and $\boldsymbol{v}$.

5. $|\boldsymbol{u} \times \boldsymbol{v}|$ gives the area of a parallelogram with two sides formed by $\boldsymbol{u}$ and $\boldsymbol{v}$:



6. $|\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})|$ represents the volume of a parallelepiped which has adjacent sides consisting of vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$. Notice that if any two of the three vectors are parallel, the result is zero.

For practice, consider $\boldsymbol{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, and $\boldsymbol{k} = (0, 0, 1)$, then using the right hand rule and the fact that these vectors are orthogonal we get:

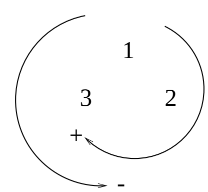
$$\boldsymbol{i} \times \boldsymbol{i} = \mathbf{0}, \qquad \boldsymbol{i} \times \mathbf{j} = \boldsymbol{k}, \qquad \boldsymbol{i} \times \boldsymbol{k} = -\mathbf{j}.$$

In indicial notation the cross product is given by

$$\boldsymbol{u} \times \boldsymbol{v} = u_i v_j \varepsilon_{ijk}$$

where ε_{ijk} is the **permutation tensor**, a third-order tensor defined as

$$\varepsilon_{ijk} = \begin{cases} 0 & \text{if any 2 indices are equal} \\ 1 & \text{if } ijk \text{ is an even permutation of 123} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123} \end{cases} \tag{1.20}$$



An even permutation of 1-2-3 is defined as taking an even number of interchanges to obtain the new order. Similarly, an odd permutation of 1-2-3 is defined as taking an odd number of interchanges to obtain the new order. So for example to determine ε_{231}, begin with 1-2-3, interchange 1 and 3 to get 3-2-1, then interchange the 3 and 2 to get 2-3-1. This required 2 interchanges (an even number) so $\varepsilon_{231} = 1$. This is a mathematical (precise) definition. In practice is is easier to use the $1 - 2 - 3$ circle depicted above. Traveling clockwise implies an even permutation and traveling counter-clockwise implies an odd permutation. To determine ε_{231}, start at 2 and then go to the 3 and then to 1. This required us to travel in the clockwise direction, so that $\varepsilon_{231} = 1$.

Exercise:

1. $\varepsilon_{123} =$ 2. $\varepsilon_{112} =$ 3. $\varepsilon_{321} =$
4. $\varepsilon_{121} =$ 5. $\varepsilon_{213} =$

Ans:
1. 1 2. 0 (two indices match) 3. -1 4. 0 (two indices match) 5. -1

It can be easily shown that the permutation tensor has the following properties:

$$\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij} \tag{1.21}$$

$$\varepsilon_{ijk} = -\varepsilon_{jik} = -\varepsilon_{ikj} = -\varepsilon_{kji} \tag{1.22}$$

So in manipulating expressions with the permutation tensor cyclical permutations do not change the sign, and interchanging 2 indices causes a sign change.

We remark that indicial notation for the cross product is generally used to manipulate vector expressions, but to compute the cross product, the determinant method, as used above, is generally more efficient.

With the permutation tensor defined, we can now express the determinant of a matrix in indicial notation. It can be shown that for a matrix $\underline{\underline{\boldsymbol{A}}}$,

$$\det[\underline{\underline{\boldsymbol{A}}}]\varepsilon_{lmn} = \varepsilon_{ijk}A_{il}A_{jm}A_{kn} \tag{1.23}$$

so that if $l = 1$, $m = 2$, $n = 3$, we have

$$\det[\underline{\underline{\boldsymbol{A}}}] = \varepsilon_{ijk}A_{i1}A_{j2}A_{k3}.$$

Multiplying both sides of equation (1.23) by ε_{lmn} we get an alternate expression,

$$6\det\underline{\underline{\boldsymbol{A}}} = \varepsilon_{lmn}\varepsilon_{ijk}A_{il}A_{jm}A_{kn}, \tag{1.24}$$

where $\varepsilon_{lmn}\varepsilon_{lmn} = \sum_l\sum_m\sum_n \varepsilon_{lmn}\varepsilon_{lmn} = \varepsilon_{111}\varepsilon_{111} + \varepsilon_{112}\varepsilon_{112} + \varepsilon_{113}\varepsilon_{113} + ... = \varepsilon_{123}\varepsilon_{123} + \varepsilon_{132}\varepsilon_{132} + \varepsilon_{213}\varepsilon_{213} + \varepsilon_{231}\varepsilon_{231} + \varepsilon_{312}\varepsilon_{312} + \varepsilon_{321}\varepsilon_{321} = 6$. To give the determinant a geometric representation (coordinate independent), recall that a tensor acting on a vector represents a mapping of the vector into another vector. So for example, if $\underline{\underline{\boldsymbol{T}}}$ is a second-order tensor which is a scalar multiple of the identity, it represents a mapping which shrinks or elongates a vector. The *definition* of the determinant for a second-order tensor in direct notation is one which gives the change of volume of a parallelepiped due to the mapping in the following sense. Recalling property 6 of the cross product regarding the volume of a parallelepiped,

Definition 1.9 *Let $\underline{\underline{\boldsymbol{T}}}$ be a second-order tensor and $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ be arbitrary vectors. Then the* **determinant** *of $\underline{\underline{\boldsymbol{T}}}$ is*

$$\left(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{u}\right)\cdot\left[(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{v})\times(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{w})\right] = \det(\underline{\underline{\boldsymbol{T}}})\left[\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})\right]. \tag{1.25}$$

We can show that our algebraic definition of determinant (the determinant of a matrix as defined in (1.23)) satisfies this relationship in the Cartesian coordinate system using indicial notation:

$$
\begin{aligned}
\left(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{u}\right)\cdot\left[(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{v})\times(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{w})\right] &= \left[T_{ij}u_j\right)\cdot\left(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{v}\times\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{w}\right]_i \\
&= (T_{ij}u_j)\cdot\left[(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{v})_k\times(\underline{\underline{\boldsymbol{T}}}\cdot\boldsymbol{w})_m\right]_i \\
&= (T_{ij}u_j)\cdot\left[T_{kl}\cdot v_l\times T_{mn}w_n\right]_i \\
&= (T_{ij}u_j)\cdot(T_{kl}v_lT_{mn}w_n\varepsilon_{kmi}) \\
&= T_{ij}T_{kl}T_{mn}\varepsilon_{kmi}u_jv_lw_n \\
&= \det(\underline{\underline{\boldsymbol{T}}})\ \varepsilon_{lnj}u_jv_lw_n \quad \text{by (1.23)} \\
&= \det(\underline{\underline{\boldsymbol{T}}})\ u_j(v_lw_n\varepsilon_{lnj}) \\
&= \det(\underline{\underline{\boldsymbol{T}}})\ u_j(v_lw_n\varepsilon_{lnk})\delta_{jk} \\
&= \det(\underline{\underline{\boldsymbol{T}}})\ \left[\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})\right].
\end{aligned}
$$

Thus in this case the algebraic and geometric definitions coincide.

We now present a very useful identity:

The Kronecker Delta - Alternating Symbol Identity ($\varepsilon - \delta$ identity):

$$\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl} \tag{1.26}$$

The order of the indices can be remembered as follows. On the left-hand side, the first indices of the two ε's must match (i in (1.26)). Let j and l be the middle indices, and k and m the last indices. Then the order of indices on the right-hand side is:
(middle-middle)(last-last) - (middle-last)(last-middle)

The proof of the $\varepsilon - \delta$ identity is left as an exercise.

Exercise:
Prove the following identity using the $\varepsilon - \delta$ identity.

$$\boldsymbol{u}\times(\boldsymbol{v}\times\boldsymbol{w}) = (\boldsymbol{u}\cdot\boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u}\cdot\boldsymbol{v})\boldsymbol{w}$$

Ans:

$$u_i\times(v_j\times w_k) = u_i\times(v_jw_k\varepsilon_{jkl}) \qquad \text{indicial notation for cross product.}$$

be sure free index introduced does not match other indices

$$= u_i v_j w_k \varepsilon_{jkl} \varepsilon_{ilm} \qquad \text{indicial notation for cross product}$$

$$= u_i v_j w_k \varepsilon_{ljk} \varepsilon_{lmi} \qquad \text{permute the indices of the permutation}$$

tensors so as to more easily apply the $\varepsilon - \delta$ identity

$$= u_i v_j w_k (\delta_{jm}\delta_{ki} - \delta_{ji}\delta_{km}) \qquad \text{apply the } \varepsilon - \delta \text{ identity}$$

$$= u_i v_j w_k \delta_{jm}\delta_{ki} - u_i v_j w_k \delta_{ji}\delta_{km} \qquad \text{note that the free index is } m$$

$$= u_i v_m w_k \delta_{ki} - u_i v_i w_k \varepsilon_{km}$$

$$= u_i v_m w_i - u_i v_i w_m \qquad \text{check that free index is still } m$$

$$= [(\boldsymbol{u}\cdot\boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u}\cdot\boldsymbol{v})\boldsymbol{w}]_m$$

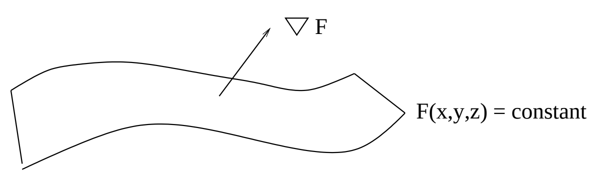
Warning: Indicial notation as presented this far is *only* applicable in an orthonormal coordinate system. This is a limitation of indicial notation as presented here. Direct notation is applicable in *any* coordinate system. Indicial notation can be extended to a non-orthonormal coordinate system, and an introduction to the formalism is presented in the last section of this chapter.

1.6 Review of Vector Calculus and More Indicial Notation

We begin by reviewing some basic vector calculus.

Recall that the **gradient** of a scalar-valued function, $F(x, y, z)$, at a point is a vector which is normal to the level surface through that point. In a Cartesian system we have

$$\boldsymbol{\nabla} F = (F_x,\ F_y,\ F_z). \tag{1.27}$$



A unit normal normal to the surface is then

$$\boldsymbol{n} = \frac{\boldsymbol{\nabla} F}{|\boldsymbol{\nabla} F|}. \tag{1.28}$$

Note that (1.27) is not coordinate independent. Hence is not a useful definition. Mathematically, the gradient is defined as follows (need REFERENCE)

Definition 1.10 *Suppose that f maps a vector field, V, into the reals. And suppose the domain of f is D and let $\boldsymbol{u} \in D$. Then the mapping $\boldsymbol{\nabla} f$ is a* gradient mapping *at $\boldsymbol{u}$ if there is a function R which maps vectors in V to the reals, i.e. $R : V \to \Re$ such that*

1. $\lim_{\boldsymbol{h}\to\boldsymbol{0}} R(\boldsymbol{h}) = \boldsymbol{0}$, *and*
2. $f(\boldsymbol{u}+\boldsymbol{h}) = f(\boldsymbol{u}) + \boldsymbol{h}\cdot\boldsymbol{\nabla} f(\boldsymbol{h}) + ||\boldsymbol{h}||R(\boldsymbol{h})$ *for any $\boldsymbol{h} \in V$ such that $\boldsymbol{u}+\boldsymbol{h} \in D$.*

This is the mathematical definition. Practically, we use the following definition:

Definition 1.11 *Let f be differentiable. Then* **the gradient of** *f, $\boldsymbol{\nabla} f$, satisfies the following relationship:*

$$df = \boldsymbol{\nabla} f \cdot d\boldsymbol{x} \tag{1.29}$$

where df is the total differential of f in space, and $d\boldsymbol{x}$ is the differential of the spatial coordinates.

The **directional derivative** of a scalar-valued function in the direction of a *unit* vector, $\boldsymbol{u}$, gives the rate of change of F in the direction of $\boldsymbol{u}$. What is desired is the magnitude of the projection of the change of F in the direction of $\boldsymbol{u}$. Using the formula for projection, (1.3), and taking the magnitude of the result, we have that the directional derivative is:

$$\boldsymbol{\nabla} F \cdot \boldsymbol{u}. \tag{1.30}$$

Suppose that the scalar-valued function $F(x, y, z)$ represents temperature, i.e. $T = F(x, y, z)$. Then $\boldsymbol{\nabla} F \cdot \boldsymbol{u}$ gives the change in temperature with respect to distance in the direction of $\boldsymbol{u}$. In what direction does the largest change occur? Since

$$\boldsymbol{\nabla} F \cdot \boldsymbol{u} = ||\boldsymbol{\nabla} F||\ ||\boldsymbol{u}|| \cos\theta,$$

$\boldsymbol{\nabla} F \cdot \boldsymbol{u}$ is maximum when $\cos\theta$ is one, or when $\boldsymbol{u}$ is in the same direction as $\boldsymbol{\nabla} F$. Thus $\boldsymbol{\nabla} F$ gives the direction of maximum (positive) change, and the magnitude of that change is given by $||\boldsymbol{\nabla} F||$. If F represents temperature, the units of $||\boldsymbol{\nabla} F||$ is degrees per unit length.

In indicial notation a *comma* denotes derivative, e.g.

$$\phi_{,i} = (\phi_{,1}, \phi_{,2}, \phi_{,3}) = \left(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}, \frac{\partial \phi}{\partial x_3} \right). \tag{1.31}$$

As before, a free index is assumed to take on the values of 1, 2, and 3, and the comma implies that the partial derivatives are taken with respect to the position coordinate. Further, one free index implies the quantity is a vector. Two free indices would imply a second-order tensor (e.g. $v_{i,j} = \boldsymbol{\nabla} \boldsymbol{v}$), etc.

Now consider a combination of derivatives with contractions:

$$v_{i,j}\, u_i = (\boldsymbol{\nabla} \boldsymbol{v} \cdot \boldsymbol{u})_j \tag{1.32}$$

$$v_{i,j}\, u_j = (\boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v})_i. \tag{1.33}$$

In (1.32) on the right-hand-side, the contraction occurs between $\boldsymbol{v}$ and $\boldsymbol{u}$. Thus the indices on $\boldsymbol{v}$ and $\boldsymbol{u}$ must match. Now consider (1.33) on the left-hand-side. The indices which match are the ones refering to the derivative of the vector $\boldsymbol{v}$ (*not* the vector itself)

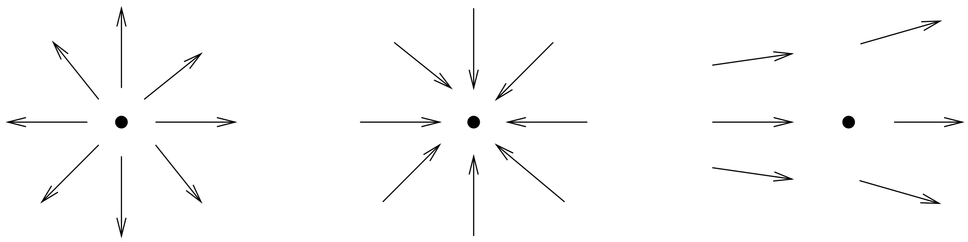


Figure 1.1: positive divergence   negative divergence   positive divergence

and the vector $\boldsymbol{u}$. Thus the contraction operator must be between the gradient and $\boldsymbol{u}$. Writing it in direct notation, tradition dictates that we place the gradient to the left of what the gradient is operating on, so that in order to contract the gradient with $\boldsymbol{u}$ we must place $\boldsymbol{u}$ to the left of the gradient.

Now consider

$$v_{i,i} = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3} = \boldsymbol{\nabla} \cdot \boldsymbol{v} \tag{1.34}$$

which is the **divergence** of the vector field $\boldsymbol{v}$. Again, notice that repeated indices implies summation or contraction. Very loosely, think that the divergence of a vector field is the amount a vector field "diverges" at a point: if $\boldsymbol{\nabla} \cdot \boldsymbol{v} > 0$, then we say the vector field is *diverging* (see Figure 1.1). Likewise, if $\boldsymbol{\nabla} \cdot \boldsymbol{v} < 0$, then the vector field is contracting (or has negative divergence).

The **curl** of a vector field roughly represents how much a vector field "swirls", and is given by:

$$(\boldsymbol{\nabla} \times \boldsymbol{v})_k = v_{j,i}\, \varepsilon_{ijk} = v_{i,j}\, \varepsilon_{jik} \tag{1.35}$$

Note that the indicial notation is identical to the cross product, except that the gradient replaces one of the vectors. The order is consistent - the index on the first vector (in this case the gradient) corresponds to the first index on the permutation tensor.

The curl is calculated similarly to the cross product. As an example consider the vector field $\boldsymbol{v} = <x^2z,\ x + 2y,\ xy^2z>$:

$$\boldsymbol{\nabla} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2z & x+2y & xy^2z \end{vmatrix}$$

$$= \boldsymbol{i}\left[\frac{\partial}{\partial y}(xy^2z) - \frac{\partial}{\partial z}(x+2y)\right] - \boldsymbol{j}\left[\frac{\partial}{\partial x}(xy^2z) - \frac{\partial}{\partial z}(x^2z)\right] + \boldsymbol{k}\left[\frac{\partial}{\partial x}(x+2y) - \frac{\partial}{\partial y}(x^2z)\right]$$

$$= <2xyz,\ x^2 - y^2z,\ 1>$$

Second and higher order derivatives in indicial notation are represented using more than one index after the comma. Consider the following examples:

1. $v_{i,jj} = \dfrac{\partial^2 v_i}{\partial x_j^2} = \dfrac{\partial}{\partial x_j}\dfrac{\partial v_i}{\partial x_j} = (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \boldsymbol{v})_i$

2. $A_{ij},_j = \dfrac{\partial A_{ij}}{\partial x_j} = (\boldsymbol{\nabla} \cdot \underline{\underline{\boldsymbol{A}}}^T)_i$

3. $v_{i},_{ij} = \dfrac{\partial^2 v_i}{\partial x_i \partial x_j} = (\boldsymbol{\nabla}\boldsymbol{\nabla} \cdot \boldsymbol{v})_j$

In the first example there are two indices following the comma indicating a second-order derivative. Since the two indices following the comma are the same there is a contraction between the derivatives, or gradients. Note that there is one free index, which means the result must be a first-order tensor, or vector.

In the second example, there is only one index following the comma so this indicates a first-order derivative. The contraction occurs between the derivative and the second component of the tensor, hence the need for the transpose on $\underline{\underline{\boldsymbol{A}}}$. Again the result is a vector since there is only one free index.

In the third example, the two indices following the comma indicate a second-order differentiation. The contraction occurs between the vector and a derivative, hence the dot product between the gradient and the vector $\boldsymbol{v}$. The result of the contraction is a scalar-valued function, so taking the gradient of this quantity makes sense and the result is again a vector.

Exercise:

Direct Notation	Indicial Notation
$\boldsymbol{w} = \boldsymbol{\nabla} \cdot \underline{\underline{\boldsymbol{T}}}$	(i)
$\boldsymbol{w} = (\underline{\underline{\boldsymbol{T}}} \cdot \boldsymbol{\nabla}) = \boldsymbol{\nabla} \cdot \underline{\underline{\boldsymbol{T}}}^T$	(ii)
(iii)	$a = T_{ij} v_{j},_i$
(iv)	$a = T_{ij} v_{i},_j$

Ans:

$(i) w_j = T_{ij},_i \quad (ii) w_i = T_{ij},_j \quad (iii) a = \underline{\underline{\boldsymbol{T}}} : \boldsymbol{\nabla}\boldsymbol{v} \quad (iv) a = \underline{\underline{\boldsymbol{T}}}^T : \boldsymbol{\nabla}\boldsymbol{v} = \underline{\underline{\boldsymbol{T}}} : (\boldsymbol{\nabla}\boldsymbol{v})^T.$

Exercise:

Show that the following identity is true in the cartesian coordinate system using indicial notation:

$$\boldsymbol{\nabla} \cdot (\boldsymbol{u} \cdot \underline{\underline{\boldsymbol{T}}}) = (\underline{\underline{\boldsymbol{T}}} \cdot \boldsymbol{\nabla}) \cdot \boldsymbol{u} + \underline{\underline{\boldsymbol{T}}} : (\boldsymbol{\nabla}\boldsymbol{u})^T \tag{1.36}$$

Ans:

$$\begin{aligned}
\boldsymbol{\nabla} \cdot (\boldsymbol{u} \cdot \underline{\underline{\boldsymbol{T}}}) &= (u_i T_{ij}),_j \\
&= u_{i},_j T_{ij} + u_i T_{ij},_j && \text{product rule} \\
&= \underline{\underline{\boldsymbol{T}}} : (\boldsymbol{u}\boldsymbol{\nabla}) + \boldsymbol{u} \cdot (\underline{\underline{\boldsymbol{T}}} \cdot \boldsymbol{\nabla}) && \text{switching into direct notation} \\
&= (\underline{\underline{\boldsymbol{T}}} \cdot \boldsymbol{\nabla}) \cdot \boldsymbol{u} + \underline{\underline{\boldsymbol{T}}} : (\boldsymbol{\nabla}\boldsymbol{u})^T && \text{switching order of terms}
\end{aligned}$$

Note that the product rule in indicial notation is applied just as it is on scalar-valued functions.

Exercise:

How would $\frac{\partial A}{\partial E_{ij}} E_{ij,k}$ be written in vector notation?

Ans:

$$\frac{\partial A}{\partial \underline{\underline{\boldsymbol{E}}}} : (\underline{\underline{\boldsymbol{E}}}\boldsymbol{\nabla})$$

1.7 Divergence Theorem, Green's Identity, and Stoke's Theorem

We now present fundamentally important theorems which are multi-dimensional forms of theorems learned in the first year of calculus. These theorems will be referred to often in this book. We begin with a definition.

Definition 1.12 *A* **regular region**, Ω, *is an open region with (i) a piecewise smooth boundary,* $\partial\Omega$, *and (ii) is such that any straight line which is parallel to any of the coordinate axes either intersects* $\partial\Omega$ *at a finite set of points or has a whole interval common with* $\partial\Omega$. *A regular region may be bounded or unbounded.*

These two criteria on the surface of Ω means that the surface can't wiggle infinitely often. This implies that the region is *orientable*, i.e. that it has a definable inside and outside, and *is differentiable almost everywhere* (from the piecewise smooth requirement) so that a unit normal outward vector may be defined almost everywhere. Mathematically the definition of a bounded regular region is an open subset of $\Re^n$ with a surface which is Lipschitz (see Thm 1.5.3.1 of Grisvard [4]). With this definition we can now state the Divergence Theorem.

Theorem 1.1 Divergence Theorem *Let* Ω *be a bounded regular region. Let each component of* $\boldsymbol{v}$ *be* $C^0(\overline{\Omega}) \cap C^1(\Omega)$, *i.e. continuous throughout the domain and its boundaries with continuous partial derivatives. Then*

$$\int_{\partial\Omega} \boldsymbol{v} \cdot \boldsymbol{n} \, d\sigma = \int_{\Omega} \boldsymbol{\nabla} \cdot \boldsymbol{v} \, dv \tag{1.37}$$

where $\boldsymbol{n}$ *is the unit outward normal vector to the surface,* $\partial\Omega$.

This allows us to convert from a volume integral in 3-dimensions of the divergence of a function into a surface integral, and vice versa.

A different form of the divergence theorem can be obtained by letting $\boldsymbol{v} = (\phi(\boldsymbol{x}), 0, 0)$, in which case we get $\int_{\partial\Omega} \phi n_1 \, d\sigma = \int_{\Omega} \frac{\partial \phi}{\partial x_1} \, dv$. Doing the same for $\boldsymbol{v} = (0, \phi(\boldsymbol{x}), 0)$ and for $\boldsymbol{v} = (0, 0, \phi(\boldsymbol{x}))$ we get

Theorem 1.2

$$\int_{\partial\Omega} \phi \boldsymbol{n} \, d\sigma = \int_{\Omega} \boldsymbol{\nabla}\phi \, dv. \tag{1.38}$$

This is just another form of the Divergence Theorem.

What does the Divergence Theorem reduce to in one dimension? Let Ω be the interval $[a, b]$. Then the boundary $\partial\Omega$ consists of two points, a and b. The unit normal to the boundary at $x = a$ is the unit vector pointing in the negative direction, i.e. $\boldsymbol{n}(x = a) = -1$. Likewise the unit normal at $x = b$ is 1. Letting $\phi(\boldsymbol{x}) = \phi(x)$ in the divergence theorem version (1.38), we get

$$\int_\Omega \boldsymbol{\nabla}\phi \, dv = \int_a^b \frac{d\phi}{dx} \, dx = \int_a^b \phi'(x) \, dx = \phi(a)\boldsymbol{n}(x = a) + \phi(b)\boldsymbol{n}(x = b) = \phi(b) - \phi(a)$$

which is the Fundamental Theorem of Calculus!

Another result easily obtained from the divergence theorem is the 3-dimensional version of integration by parts. Begin with (1.38) and let $\phi = fg$ where f and g are scalar valued functions. We then obtain

$$\int_{\partial\Omega} fg\boldsymbol{n} \, d\sigma = \int_\Omega (f\boldsymbol{\nabla}g + g\boldsymbol{\nabla}f) \, dv$$

where we have used $\boldsymbol{\nabla}(fg) = f\boldsymbol{\nabla}g + g\boldsymbol{\nabla}f$ which can be shown using indicial notation by using the product rule. Rearranging terms yields

Green's Identity

$$\int_\Omega f\boldsymbol{\nabla}g \, dv = \int_{\partial\Omega} fg\boldsymbol{n} \, d\sigma - \int_\Omega g\boldsymbol{\nabla}f \, dv. \tag{1.39}$$

This is exactly the multi-dimensional form of integration by parts for one-dimension, $\int_a^b u \, dv = uv|_a^b - \int_a^b v \, du$. Recall that in proving Green's identity, the divergence theorem is used, so that all restrictions placed on Ω in the divergence theorem apply here as well.

Stoke's Theorem is an integral relation involving the curl of a vector field instead of the divergence. It relates a vector field defined on a surface which exists in 3-dimensions, S, and the vector field defined on the curve which makes up the boundary of S, C. Think of the surface as being a hat and C as being the outer brim of the hat. Mathematically, S is called a 2-dimensional manifold. In order for Stoke's theorem to hold, S must be a surface which is piecewise smooth (if the surface is parametrized by $z = (x(t), \; y(t))$, then $x(t)$ and $y(t)$ must be differentiable almost everywhere) and *orientable*, i.e. has 2 sides, one with a defined normal $\boldsymbol{n}$ and the other side with a normal $-\boldsymbol{n}$ (almost everywhere). Since the surface is piecewise smooth, the normal is defined almost everywhere. An example of a non-orientable surface is the mobius strip: take a strip of paper, give it a half-twist so that one side which was originally up is now down, then join the two (far) ends as if making a ring. The mobius strip has niether an inside nor an outside, and is hence not "orientable". We assume the curve is oriented in such a way that following the curve with the right-hand fingers orients the thumb in the same direction as the unit normal, $\boldsymbol{n}$ (see figure 1.2).

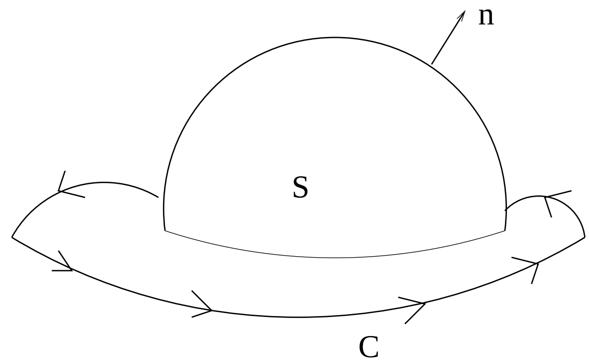


Figure 1.2: An example of a surface with the curve oriented appropriately for the normal illustrated.

Theorem 1.3 Stoke's Theorem
Given an orientable surface, S, bounded by a closed, regular curve C, then

$$\int\int_S \boldsymbol{n} \cdot (\boldsymbol{\nabla} \times \boldsymbol{v})\, ds = \int_C \boldsymbol{v} \cdot d\boldsymbol{l} \tag{1.40}$$

where $\boldsymbol{v}$ is a $C^1(S) \cap C^0(\overline{S})$ vector field.

.

Recall that if C is parametrized by $\boldsymbol{r}(t) = (x(t),\ y(t),\ z(t))$, in the proper orientation, then the line integral becomes $\int_{t=a}^{t=b} \boldsymbol{v}(t) \cdot \boldsymbol{r}'\, dt$.

1.8 Fourth-Order Tensors with Minor Symmetries

A fourth-order tensor is a linear operator acting on lower order tensors. In this section we will concentrate on fourth-order tensors acting on symmetric second-order tensors which frequently appears in solid mechanics. A typical expression is

$$\sigma_{ij} = E_{ijkl} e_{kl} \tag{1.41}$$

where $\underline{\underline{\boldsymbol{\sigma}}}$ and $\underline{\underline{\boldsymbol{e}}}$ are symmetric second-order tensors. In solid mechanics, $\underline{\underline{\boldsymbol{\sigma}}}$ is called the *stress tensor*, $\underline{\underline{\boldsymbol{e}}}$ is the strain tensor, and $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ is the stiffness tensor which gives an idea of how difficult it is to deform the solid. Because of the symmetry in $\underline{\underline{\boldsymbol{e}}}$ and $\underline{\underline{\boldsymbol{\sigma}}}$, $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ has *minor symmetries*:

$$E_{ijkl} = E_{jikl} = E_{ijlk}, \tag{1.42}$$

i.e. the first two and last two indices can be interchanged. Sometimes a fourth-order tensor has a *major* symmetry, and this is given by: $E_{ijkl} = E_{klij}$, but this is not assumed for $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ in this section.

1.8.1 Identity

The fourth-order identity tensor is defined so that

$$\underline{\underline{\underline{\underline{\boldsymbol{I}}}}} : \underline{\underline{\boldsymbol{A}}} = \underline{\underline{\boldsymbol{A}}} \tag{1.43}$$

$$\underline{\underline{\underline{\underline{\boldsymbol{I}}}}} : \underline{\underline{\underline{\underline{\boldsymbol{E}}}}} = \underline{\underline{\underline{\underline{\boldsymbol{E}}}}}, \tag{1.44}$$

i.e. contracting a fourth-order tensor with either a second-order or fourth-order tensor preserves the tensor. One way to define the fourth-order might be, $I_{ijkl} = \delta_{ij}\delta_{kl}$ but this does not preserve a fourth-order tensor: $I_{ijkl}E_{klmn} = E_{kkmn}\delta_{ij}$. Another possible definition might be

$$I_{ijkl} = \delta_{ik}\delta_{jl} \tag{1.45}$$

so that $I_{ijkl}E_{klmn} = E_{ijmn}$, however this definition of $\underline{\underline{\underline{\underline{\boldsymbol{I}}}}}$ does not have minor symmetry (something one would expect of the identity tensor and which, if the identity does not have this property, could cause problems). So we define explicitly a symmetric fourth-order identity tensor:

$$I^s_{ijkl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{1.46}$$

This definition has both minor and major symmetries, and satisfies (1.43) and (1.44) provided $\underline{\underline{\boldsymbol{A}}}$ is symmetric and $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ has a minor symmetry in the first two indices.

With this definition we can define the inverse of a fourth-order tensor: The inverse of a fourth-order tensor, $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$, is the tensor $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}^{-1}$ such that $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}} : \underline{\underline{\underline{\underline{\boldsymbol{E}}}}}^{-1} = \underline{\underline{\underline{\underline{\boldsymbol{E}}}}}^{-1} : \underline{\underline{\underline{\underline{\boldsymbol{E}}}}} = \underline{\underline{\underline{\underline{\boldsymbol{I}}}}}^s$. Just as is the case for second-order tensors, not all fourth-order tensors have inverses, and as long as one uses this definition of inverse, expected properties, such as uniqueness of the inverse, follow.

1.8.2 6×6 Component Notation of Fourth-Order Tensors

In general it's not easy to work with fourth-order tensors in its generality - but for a fourth-order tensor with minor symmetries (such as $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$) it can be represented as a 6×6 matrix. Begin with the symmetric tensors $\underline{\underline{\boldsymbol{\sigma}}}$ and $\underline{\underline{\boldsymbol{e}}}$ which, due to symmetry, can be represented as 1×6 vectors:

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{bmatrix}, \qquad \boldsymbol{e} = \begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ e_{12} \\ e_{23} \\ e_{31} \end{bmatrix}. \tag{1.47}$$

It turns out that this is not a good representation for several reasons. The first is seen by calculating

$$\underline{\underline{\boldsymbol{\sigma}}} : \underline{\underline{\boldsymbol{e}}} = \sigma_{ij}e_{ij} \quad = \quad \sigma_{11}e_{11} + \sigma_{12}e_{12} + \sigma_{13}e_{13} \tag{1.48}$$

$$
\begin{aligned}
&+\sigma_{21}e_{21} + \sigma_{22}e_{22} + \sigma_{23}e_{23} \\
&+\sigma_{31}e_{31} + \sigma_{32}e_{32} + \sigma_{33}e_{23} \\
= \;& \sigma_{11}e_{11} + \sigma_{22}e_{22} + \sigma_{33}e_{33} \\
&+2\sigma_{12}e_{12} + 2\sigma_{23}e_{23} + 2\sigma_{31}e_{31}
\end{aligned} \tag{1.49}
$$

where we used the symmetry of $\underline{\underline{\boldsymbol{\sigma}}}$ and $\underline{\underline{\boldsymbol{e}}}$ for the second equality. Now we can see that the vector notation of $\underline{\underline{\boldsymbol{\sigma}}}$ and $\underline{\underline{\boldsymbol{e}}}$ as given in (1.47) does not give: $\boldsymbol{\sigma} \cdot \boldsymbol{e} = \underline{\underline{\boldsymbol{\sigma}}} : \underline{\underline{\boldsymbol{e}}}$. To rectify this situation (and others) the definition of the 1×6 vector representations are modified, with the most common being the *Voigt notation*:

$$
\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{bmatrix}, \qquad \boldsymbol{e} = \begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ 2e_{12} \\ 2e_{23} \\ 2e_{31} \end{bmatrix}. \tag{1.50}
$$

Using the Voigt notation we now have $\boldsymbol{\sigma} \cdot \boldsymbol{e} = \underline{\underline{\boldsymbol{\sigma}}} : \underline{\underline{\boldsymbol{e}}}$. Another option which also preserves this property is the *Kelvin notation*:

$$
\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\sigma_{12} \\ \sqrt{2}\sigma_{23} \\ \sqrt{2}\sigma_{31} \end{bmatrix}, \qquad \boldsymbol{e} = \begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ \sqrt{2}e_{12} \\ \sqrt{2}e_{23} \\ \sqrt{2}e_{31} \end{bmatrix}. \tag{1.51}
$$

Althought the Kelvin notation is more mathematically pleasing (not favoring one tensor over the other), it is not as common as the Voigt notation. As a side note, the Voigt notation for $\underline{\underline{\boldsymbol{\sigma}}}$ and $\underline{\underline{\boldsymbol{e}}}$ are not tensors in the engineering sense.

The next step is to determine how $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ is represented using the Voigt (and as it turns out, the Kelvin) notation. Checking by writing out each term of $\underline{\underline{\boldsymbol{\sigma}}} = \underline{\underline{\underline{\underline{\boldsymbol{E}}}}} : \underline{\underline{\boldsymbol{e}}}$, one can check that

$$
\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{bmatrix} = \left[\begin{array}{ccc|ccc} E_{1111} & E_{1122} & E_{1133} & E_{1112} & E_{1123} & E_{1131} \\ E_{2211} & E_{2222} & E_{2233} & E_{2212} & E_{2223} & E_{2231} \\ E_{3311} & E_{3322} & E_{3333} & E_{3312} & E_{3323} & E_{3331} \\ \hline E_{1211} & E_{1222} & E_{1233} & E_{1212} & E_{1223} & E_{1231} \\ E_{2311} & E_{2322} & E_{2333} & E_{2312} & E_{2323} & E_{2331} \\ E_{3111} & E_{3122} & E_{3133} & E_{3112} & E_{3123} & E_{3131} \end{array}\right] \begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ 2e_{12} \\ 2e_{23} \\ 2e_{31} \end{bmatrix}. \tag{1.52}
$$

Thus if $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ has major symmetry then the 6×6 notation of $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$ is also symmetric. Note that if we used the notation given by (1.47) there would be some scalar factors of 2 floating in the matrix.

Now let's consider the inverse of $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}}$, $\underline{\underline{\underline{\underline{\boldsymbol{C}}}}}$, so that $\underline{\underline{\underline{\underline{\boldsymbol{E}}}}} : \underline{\underline{\underline{\underline{\boldsymbol{C}}}}} = \underline{\underline{\underline{\underline{\boldsymbol{I}}}}}^s$, and $\underline{\underline{\boldsymbol{e}}} = \underline{\underline{\underline{\underline{\boldsymbol{C}}}}} : \underline{\underline{\boldsymbol{\sigma}}}$. Using the Voigt notation we get

$$\begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ 2e_{12} \\ 2e_{23} \\ 2e_{31} \end{bmatrix} = \left[\begin{array}{ccc|ccc} C_{1111} & C_{1122} & C_{1133} & 2C_{1112} & 2C_{1123} & 2C_{1131} \\ C_{2211} & C_{2222} & C_{2233} & 2C_{2212} & 2C_{2223} & 2C_{2231} \\ C_{3311} & C_{3322} & C_{3333} & 2C_{3312} & 2C_{3323} & 2C_{3331} \\ \hline 2C_{1211} & 2C_{1222} & 2C_{1233} & 4C_{1212} & 4C_{1223} & 4C_{1231} \\ 2C_{2311} & 2C_{2322} & 2C_{2333} & 4C_{2312} & 4C_{2323} & 4C_{2331} \\ 2C_{3111} & 2C_{3122} & 2C_{3133} & 4C_{3112} & 4C_{3123} & 4C_{3131} \end{array}\right] \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{bmatrix} . \qquad (1.53)$$

Because of the choice of putting the factor of 2 into $\underline{\underline{\boldsymbol{e}}}$ the coefficients 2 and 4 appear in the 6×6 matrix representation of $\underline{\underline{\underline{\underline{\boldsymbol{C}}}}}$. If the Kelvin notation is used, there are no coefficients which appear, and if neither the Kelvin nor Voigt notation is used, the 6×6 matrix representation of $\underline{\underline{\underline{\underline{\boldsymbol{C}}}}}$ is not symmetric even if $\underline{\underline{\underline{\underline{\boldsymbol{C}}}}}$ has major symmetry (not checked).

In short, one should be very careful using a 6×6 matrix to represent a fourth-order tensor with minor symmetries. If unsure, it is best to check the representation against the expansion of the indicial notation.

1.9 Introduction to Indicial Notation in a Curvilinear Coordinate System

(from Rebecca Brannon's notes)

Up to this point, indicial notation has been presented assuming we are using the Cartesian, or rectangular, coordinate system. However many times another coordinate system, such as the cylinder or spherical coordinate system, is the more natural system to use.

In order to introduce basis into indicial notation we begin with the Cartesian system. Let $\{\boldsymbol{e}^1,\ \boldsymbol{e}^2,\ \boldsymbol{e}^3\}$ represent the standard basis, i.e. $\boldsymbol{e}^1 = [1,0,0]^T$ $\boldsymbol{e}^2 = [0,1,0]^T$ $\boldsymbol{e}^3 = [0,0,1]^T$. Then a vector and second-order tensor are represented completely as:

$$\boldsymbol{v} = v_i \boldsymbol{e}_i \qquad (1.54)$$

$$\underline{\underline{\boldsymbol{T}}} = T_{ij} \boldsymbol{e}^i \otimes\ \boldsymbol{e}^j \qquad (1.55)$$

where $\otimes$ is the tensor product (see (1.17) for definition). It is now possible to distinguish between a tensor and its transpose:

$$\underline{\underline{\boldsymbol{B}}} = B_{ij} \boldsymbol{e}^i \otimes \boldsymbol{e}^j \qquad (1.56)$$

$$\underline{\underline{\boldsymbol{B}}}^T = B_{ij} \boldsymbol{e}^j \otimes \boldsymbol{e}^i. \qquad (1.57)$$

Exercise: Write the following expressions in complete indicial notation, including basis:

1. $\boldsymbol{\nabla v}$

2. $(\boldsymbol{\nabla v})^T$
3. $\boldsymbol{uv}$
4. $\underline{\underline{\boldsymbol{A}}} : \underline{\underline{\boldsymbol{B}}}$

Ans: (i) $v_{i,j}\, \boldsymbol{e}^j \otimes \boldsymbol{e}^i$ (ii) $v_{i,j}\, \boldsymbol{e}^i \otimes \boldsymbol{e}^j$ (iii) $u_i v_j \boldsymbol{e}^i \otimes \boldsymbol{e}^j$
(iv)

$$\begin{aligned} \underline{\underline{\boldsymbol{A}}} : \underline{\underline{\boldsymbol{B}}} &= (A_{ij}\boldsymbol{e}^i \otimes \boldsymbol{e}^j) : (B_{kl}\boldsymbol{e}^k \otimes \boldsymbol{e}^l) \\ &= A_{ij}B_{kl}(\boldsymbol{e}^i \otimes \boldsymbol{e}^j) : (\boldsymbol{e}^k \otimes \boldsymbol{e}^l) \\ &= A_{ij}B_{kl}(\boldsymbol{e}^i \cdot \boldsymbol{e}^k)(\boldsymbol{e}^j \cdot \boldsymbol{e}^l) \\ &= A_{ij}B_{kl}\delta_{ik}\delta_{jl} = A_{ij}B_{ij} \end{aligned}$$

so that the end result is a scalar and no basis is needed.

Exercise: Write the following expressions in direct notation.

1. $A_{ij}B_{jk}\boldsymbol{e}^i \otimes \boldsymbol{e}^k$
2. $A_{ij}B_{kj}\boldsymbol{e}^i \otimes \boldsymbol{e}^k$
3. $A_{ji}B_{kj}\boldsymbol{e}^i \otimes \boldsymbol{e}^k$
4. $A_{ji}B_{jk}\boldsymbol{e}^i \otimes \boldsymbol{e}^k$
5. $A_{ji}B_{jk}\boldsymbol{e}^k \otimes \boldsymbol{e}^i$

Ans: With the basis vectors given there is no ambiguity regarding the transpose of the resulting matrix (second-order tensor). The index on the first basis vector ($\boldsymbol{e}^i$ for the first 4 cases) tells us that the term with the i indice must appear on the left. Thus in (iii) we know that the in direct notation we must have $\underline{\underline{\boldsymbol{A}}}^T$ and not $\underline{\underline{\boldsymbol{A}}}$. (i) $\underline{\underline{\boldsymbol{A}}} \cdot \underline{\underline{\boldsymbol{B}}}$ (ii) $\underline{\underline{\boldsymbol{A}}} \cdot \underline{\underline{\boldsymbol{B}}}^T$ (iii) $\underline{\underline{\boldsymbol{A}}}^T \cdot \underline{\underline{\boldsymbol{B}}}^T$ (iv) $\underline{\underline{\boldsymbol{A}}}^T \cdot \underline{\underline{\boldsymbol{B}}}$ (v) $\underline{\underline{\boldsymbol{B}}}^T \cdot \underline{\underline{\boldsymbol{A}}}$ which is the transpose of (iv).

We next formerly introduce cylindrical and spherical coordinates. In cylindrical coordinates, the **physical coordinates** are given by (r, θ, z) and using the appropriate basis we have any vector may be represented as

$$\boldsymbol{v} = r\boldsymbol{e}_r + \theta\boldsymbol{e}_\theta + z\boldsymbol{e}_z. \tag{1.58}$$

In spherical cooridinates, the **physical coordinates** are given by (ρ, ϕ, θ) where ϕ is the angle from the z-axis and θ is the angle from the x-axis in the xy-plane (see Figure). The order of the coordinates are important as this order gives a right-hand coordinate system whereas (ρ, θ, ϕ) does not.

Both of these basis are **orthogonal basis** since the base vectors are mutually orthogonal. However, they are *not* **homogeneous basis** since the base vectors change direction depending on where they are in space. Non-homogeneous basis are termed

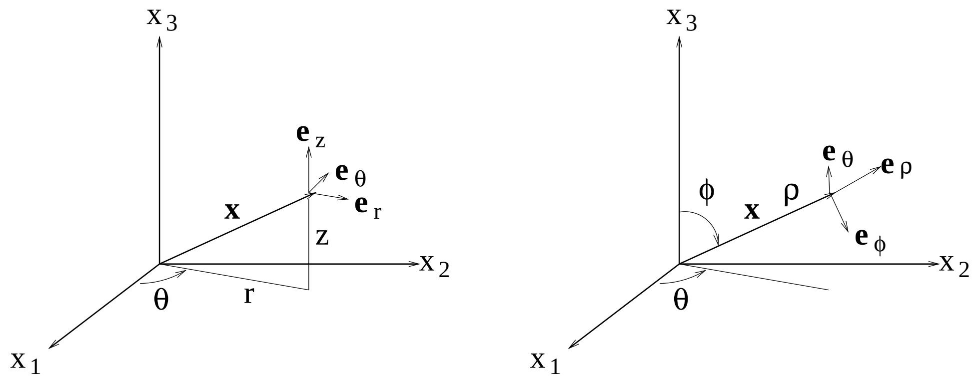


Figure 1.3: Cylindrical Coordinates Spherical Coordinates

curvilinear basis. Thus both cylindrical and spherical coordinate systems are curvilinear orthorgonal coordinate systems. In contrast, the Cartesian coordinate system is a homogeneous orthogonal coordinate system. There are systems which are both curvilinear and non-orthogonal, or homogeneous non-orthogonal.

By definition, the **covariant base vectors**, denoted as $\{\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3\}$, must point in the direction that the position vector $\boldsymbol{x}$ moves when changing the associated coordinate holding the other position coordinates constant. By convention the coordinate associated with basis vector $\boldsymbol{g}_i$ is denoted by v^i with the index now in the superscript position. Thus $\boldsymbol{v} = v^i \boldsymbol{g}_i$ where summation is implied and i is a dummy variable. So for example, in cylindrical coordinates, $\boldsymbol{g}_r$ points in the direction which is obtained if one fixes θ and z and increases r, i.e. away from the x_3-axis and parallel to the x_1x_2-plane (see Figure 1.3). The base vectors do not have to be of unit length.

Next consider the dot product:

$$
\begin{aligned}
\boldsymbol{v} \cdot \boldsymbol{v} &= (v^i \boldsymbol{g}_i) \cdot (v^j \boldsymbol{g}_j) \\
&= v^i v^j \boldsymbol{g}_i \cdot \boldsymbol{g}_j \\
&= v^1 v^1 \boldsymbol{g}_1 \cdot \boldsymbol{g}_1 + v^1 v^2 \boldsymbol{g}_1 \cdot \boldsymbol{g}_2 + v^1 v^3 \boldsymbol{g}_1 \cdot \boldsymbol{g}_3 \\
&\quad + v^2 v^1 \boldsymbol{g}_2 \cdot \boldsymbol{g}_1 + v^2 v^2 \boldsymbol{g}_2 \cdot \boldsymbol{g}_2 + v^2 v^3 \boldsymbol{g}_2 \cdot \boldsymbol{g}_3 \\
&\quad + v^3 v^1 \boldsymbol{g}_3 \cdot \boldsymbol{g}_1 + v^3 v^2 \boldsymbol{g}_3 \cdot \boldsymbol{g}_2 + v^3 v^3 \boldsymbol{g}_3 \cdot \boldsymbol{g}_3.
\end{aligned}
$$

Because we are in a more general coordinate system, *there is no guarantee that* $\boldsymbol{g}_i \cdot \boldsymbol{g}_j = \delta_{ij}$. To remedy this situation, we introduce a *contravariant basis*, $\boldsymbol{g}^i$ so that $\boldsymbol{g}_i \cdot \boldsymbol{g}^j = \delta_i^j$ where

$$
\delta_i^j = \begin{cases} 1 & \text{if } \; i = j \\ 0 & \text{if } \; i \neq j \end{cases}
$$

is the Kronecker delta. By convention summation is implied if a repeated index occurs once in the subscript position and once in the superscript position. Thus, e.g. $\delta_i^i = 3$.

1.10 Exercises

1. (3 pts) Let

$$\underline{\underline{\boldsymbol{A}}} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} \qquad \underline{\underline{\boldsymbol{B}}} = \begin{bmatrix} -1 & -2 & -3 \\ -4 & -5 & 4 \\ 3 & 2 & 1 \end{bmatrix}.$$

   Calculate $\underline{\underline{\boldsymbol{A}}} : \underline{\underline{\boldsymbol{B}}}$.

2. (4 pts) Express $b_k = u_i v_j \varepsilon_{ijk}$ in terms of (b_1, b_2, b_3), (u_1, u_2, u_3), and (v_1, v_2, v_3).

3. (24 pts) Simplify the following expressions:

   (a) $u_i v_j \delta_{ij}$

   (b) $v_k \delta_{2j} \delta_{jk}$

   (c) $\varepsilon_{i3k} \delta_{ip} v_k$

   (d) δ_{ii}

   (e) $\delta_{ij} \delta_{ij}$

   (f) $\varepsilon_{ijk} \varepsilon_{jki}$

   (g) $\delta_{ij} \delta_{jk}$

   (h) $\delta_{ij} \varepsilon_{ijk}$

4. (9 pts) What is wrong with the following indicial equations? Correct the expressions.
   (a) $w_i = b_{ik} u_i v_k$ (b) $\phi = b_{ik} u_i$ (c) $\phi_{jp} = R_{ijkl} T_{kl} u_p$

5. (7 pts) Show that the relation $\boldsymbol{v} = (\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n} + \boldsymbol{n} \times (\boldsymbol{v} \times \boldsymbol{n})$ holds for all nonzero unit vectors $\boldsymbol{n}$ and that this represents a resolution of $\boldsymbol{v}$ into vectors parallel and perpendicular to $\boldsymbol{n}$. Use the $\delta - \varepsilon$ identity.

6. (7 pts) (from R. Brannon) Use the $\delta - \varepsilon$ identity to determine what the question marks stand for: $(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = (\beta? \cdot \beta?)\beta? - (\beta? \cdot \beta?)\beta?$ where $\beta?$ are all vectors. Explain why $\boldsymbol{a} \times \boldsymbol{b} \times \boldsymbol{c}$ is poor notation because it is ambiguous without parentheses. Explain why $\boldsymbol{n} \times \boldsymbol{u} \times \boldsymbol{n}$ is acceptable because it is *not* ambiguous.

7. (9 pts) Prove the $\delta - \varepsilon$ identity $\varepsilon_{ijk} \varepsilon_{irs} = \delta_{jr} \delta_{ks} - \delta_{js} \delta_{kr}$. One way is to construct a table with various combinations of indices. Alternatively, Malvern (page 25) provides the following hint. Two of the four free indices must be equal. Show that if $j = k$ or $r = s$, then both sides vanish. Then show that even with $j \neq k$ and $r \neq s$ both sides vanish when $j = r$ unless also $k = s$. What other cases must be considered?

8. (8 pts) (Malvern, Page 46, 11 and 12). Use indicial notation, and do not assume any other results.

(a) Show that $\text{tr}(\underline{\underline{\boldsymbol{T}}} \cdot \underline{\underline{\boldsymbol{U}}}) = \text{tr}(\underline{\underline{\boldsymbol{T}}}^{\mathrm{T}} \cdot \underline{\underline{\boldsymbol{U}}})$ if either $\underline{\underline{\boldsymbol{T}}}$ or $\underline{\underline{\boldsymbol{U}}}$ is symmetric.

(b) Show that $\text{tr}(\underline{\underline{\boldsymbol{T}}} \cdot \underline{\underline{\boldsymbol{U}}}) = 0$ if one of the tensors is anti-symmetric (or skew-symmetric) and the other is symmetric.

9. (15 pts) Write in indicial notation:

   (a) $\boldsymbol{\nabla}\phi$

   (b) curl$\boldsymbol{v}$

   (c) $\nabla^2\boldsymbol{v} = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\boldsymbol{v} \equiv \triangle\boldsymbol{v}$

   (d) $\underline{\underline{\boldsymbol{T}}} \otimes \boldsymbol{v}$

   (e) $\underline{\underline{\boldsymbol{T}}} \cdot \boldsymbol{v}$

10. (20 pts) (Malvern pg 56). Assume F is a scalar-valued function and $\boldsymbol{v}$, $\boldsymbol{u}$ are vectors which are smooth enough to take the appropriate number of derivatives. Use indicial notation to prove the following identities in cartesian coordinates.

   (a) $\boldsymbol{\nabla} \cdot (F\boldsymbol{v}) = (\boldsymbol{\nabla}F) \cdot \boldsymbol{v} + F\boldsymbol{\nabla} \cdot \boldsymbol{v}$

   (b) $\boldsymbol{\nabla} \times (\boldsymbol{\nabla}F) = \boldsymbol{0}$

   (c) $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{v}) = 0$

   (d) $\boldsymbol{\nabla} \cdot (\boldsymbol{u} \times \boldsymbol{v}) = (\boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{v} - \boldsymbol{u} \cdot (\boldsymbol{\nabla} \times \boldsymbol{v})$

   (e) $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{v}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{v}) - \boldsymbol{\nabla}^2\boldsymbol{v}$

11. (6 pts) Write out explicitly (three equations with all indices given as 1, 2, or 3) in terms of ρ, $\boldsymbol{v}$, p, $\boldsymbol{g}$, λ, μ:

$$\rho\frac{D\boldsymbol{v}}{Dt} + \boldsymbol{\nabla}p - \boldsymbol{\nabla}(\lambda\boldsymbol{\nabla} \cdot \boldsymbol{v}) - \boldsymbol{\nabla} \cdot (2\mu\underline{\underline{\boldsymbol{d}}}) = \rho\boldsymbol{g} \tag{1.59}$$

where $\underline{\underline{\boldsymbol{d}}} = \frac{1}{2}[\boldsymbol{\nabla}\boldsymbol{v} + (\boldsymbol{\nabla}\boldsymbol{v})^T]$. This is the Navier-Stokes equation which describes the flow of fluid. Here ρ is the density of the fluid, $\boldsymbol{v}$ is the velocity, p is pressure, $\boldsymbol{g}$ is an externally applied force (gravity), and λ and μ are constants.

12. (8 pts) Let σ_{ij} and e_{ij} be symmetric tensors, and C_{ijkl} be a 4th-order tensor given by
$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) + k(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$.
where $\sigma_{ij} = C_{ijkl}e_{kl}$.

   (a) Show that $\sigma_{12} = 2\mu e_{12}$.

   (b) Write σ_{ij} as a vector: $[\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{13}, \sigma_{23}]$, and e_{ij} as a vector: $[e_{11}, e_{22}, e_{33}, e_{12}, e_{13}, e_{23}]$. Find the 2-dimensional tensor (matrix) representing C_{ijkl} such that $\sigma_i = C'_{ij}e_j$ in terms of λ and μ. This is a technique which can be used to view a 4th-order tensor.

   (c) Normally a 4th order tensor has 81 components. Why can we represent C_{ijkl} as a matrix with only 36 components?

13. (6 pts) (Malvern, Pg. 61, No. 2). A force of magnitude F acts in a direction radially away from the origin at a point $(\frac{a}{3}, \frac{2b}{3}, \frac{2c}{3})$ on the surface of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Determine the component of the force vector in the direction normal to the surface.

14. (6 pts) (i) If $\boldsymbol{r} = (x_1, x_2, x_3)$ is the position vector, use the divergence theorem to express $\int_{\partial R} \boldsymbol{r} \cdot \boldsymbol{n} \, ds$ in terms of the volume of the region R. (ii) Actually perform the surface integral for a unit cube with one corner at the origin of a Euclidean coordinate system.

15. (7 pts) (Malvern, Pg. 212, No. 2) A planar area in the x_1x_2-plane is bounded by the square with corners $(0,0)$, $(b,0)$, (b,b), $(0,b)$. A vector $\boldsymbol{v}$ has components $v_1 = Ax_2$, $v_2 = Bx_2$, $v_3 = 0$, where A and B are constants.

    (a) Verify that the divergence theorem holds for a cube with one surface coincident with the plane area.

    (b) Verify that Stokes' theorem holds on the x_1x_2 planar area.