CHAPTER 1

THE MATHEMATICS OF ELECTRICITY AND MAGNETISM

As one proceeds further into the study of physics, mathematics becomes a more important tool in physical theory. Without a background in mathematical concepts we are limited in how we can describe, use and extend the results which we obtain from experiments. Thus in this first chapter we will introduce the important definitions and concepts of those topics in mathematics which are particularly necessary to describe the basic theories of electricity and magnetism. It turns out that the mathematics of vectors is particularly important because the fundamental experimental results which lead to the classical theory of electricity and magnetism can be expressed in terms of a few vector equations.

Although we will emphasize vectors in our mathematical introduction, we will also introduce the other mathematical concepts, tensors and complex variables for example, which are useful in describing specific aspects of electromagnetic theory. We will also study the properties of different coordinate systems.

It is not our intent to develop a comprehensive study of these mathematical topics. We wish only to introduce common mathematical techniques which are found to be particularly useful in electricity and magnetism. The reader is referred to the references at the end of the chapter for a more thorough discussion of this material. Students who have adequate knowledge of the mathematics can move directly to Chapter II where we begin the study of electricity.

1-1: Scalars, Vectors and Vector Addition

In classical physics many physical quantities can be represented by the mathematical quantities of scalars or vectors. It is important to understand the differences between these quantities.

A ***scalar*** describes a physical quantity which can be completely defined by a single number only such as mass, time, temperature, electrical charge and potential energy. Scalar quantities obey the algebraic relationships of ordinary numbers. For example if a, b, and c are scalars then:

$$a + b = b + a \quad \text{(commutative law of addition)} \qquad (1\text{-}1)$$

$$a + (b + c) = (a + b) + c \quad \text{(associative law of addition)} \qquad (1\text{-}2)$$

$$ab = ba \quad \text{(commutative law of multiplication)} \qquad (1\text{-}3)$$

$$a(bc) = (ab)c \quad \text{(associative law of multiplication)} \qquad (1\text{-}4)$$

$$a(b + c) = ab + ac \quad \text{(distributive law)} \qquad (1\text{-}5)$$

A ***vector*** quantity differs from a scalar quantity in that it must be defined by two quantities, one representing the magnitude of the quantity and the other representing the direction in which the quantity acts. Some examples of physical quantities which require representation by vectors are force, velocity (as distinguished from speed), and displacement (as distinguished from length). In the case of force, for example, we know that if it is acting on an object we have no way of estimating its effect upon that object unless we know both the magnitude and the direction in which the force is acting. In this text vectors are represented by bold face symbols (i.e. **A**, $\boldsymbol{\omega}$, $\boldsymbol{\nabla}$).

Pictorially a vector may be represented by an arrow pointing in the direction of action of the vector and with a length proportional to the magnitude of the vector. In Figure 1-1 we show how this can be done. The vectors **A** and **B** are in the same

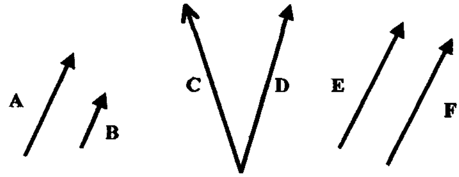


Figure 1-1: Representation of Vectors

direction but of different magnitude. They are therefore not equal. The vectors **C** and **D** are of the same magnitude from the same point but again are not equal because they are in different directions. The vectors **E** and **F** are of the same magnitude and lie in the same direction. They are equal vectors, even though they do not share a common point.

Vectors may be combined through addition. A geometric illustration of the addition of vectors may be made by use of the parallelogram rule which we illustrate in Figure 1-2. Given two vectors **A** and **B** the sum, which we call **C**, is given by

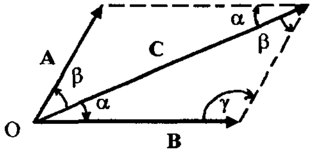


Figure 1-2: Diagram for Vector Addition

determining the diagonal of the parallelogram constructed as shown in Figure 1-2. We draw one side parallel to the direction of **A** and of length equal to the magnitude of **A**, which we denote as |**A**| or simply A. We draw the other side parallel to the direction of **B** and of length equal to the magnitude, |**B**|. We should note that this definition is more than a mathematical construct. We know that physical quantities represented by vectors combine in this way and our mathematical definition is intended to describe physical reality. For example, if **A** and **B** in Figure 1-2 represent forces acting on a point object at the point O in the figure, then **C** will represent the magnitude and direction of the resultant force acting on the object, because forces act as vectors.

We find the magnitude and direction of **C** from the parameters of **A** and **B**. The magnitude of the vector is found from the cosine law:

$$|\mathbf{C}|^2 = |\mathbf{A}|^2 + |\mathbf{B}|^2 - 2\,|\mathbf{A}||\mathbf{B}|\cos\gamma = A^2 + B^2 - 2AB\cos\gamma$$
$$= A^2 + B^2 + 2AB\cos(\alpha+\beta) \quad (1\text{-}6)$$

where we have used the fact that $\alpha + \beta = \pi - \gamma$. This latter expression is usually the most convenient one to use because it involves the angle between the vectors **A** and **B** which is often known.

The direction of the resultant vector is found from the sine law:

$$\frac{A}{\sin\alpha} = \frac{B}{\sin\beta} = \frac{C}{\sin\gamma} \quad (1\text{-}7)$$

thus, since we want α,

$$\sin\alpha = \sin\beta\,\tfrac{A}{B} = \sin\gamma\,\tfrac{A}{C} = \sin(\alpha+\beta)\,\tfrac{A}{C} \quad (1\text{-}8)$$

and the parameters of the vector **C** are determined.

In the actual situation of working with vectors most often the vectors will be referenced to some specific coordinate system. The vector thus can be broken down into components which relate to the specific coordinate system used. If, for example, we have a vector **A** which is defined in the Cartesian coordinate system it can be represented as a sum of three vectors, $\mathbf{A}_X$, $\mathbf{A}_Y$, and $\mathbf{A}_Z$ which are the components of **A** along the x, y and z axes as we have illustrated in Figure 1-3.

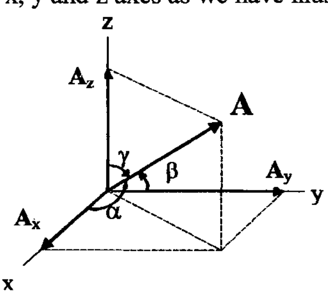


Figure 1-3: Cartesian Coordinate Representation of a Vector

We will have cause to use what are called ***unit vectors***. These are vectors of magnitude one which we represent by letters in bold face type and a circumflex such as $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$ or $\hat{\mathbf{e}}_z$, $\hat{\mathbf{e}}_\phi$, etc., and which we define according to the coordinate system being used. The advantage of using unit vectors is that it is a very simple way to separate the magnitude of the vector from its direction.

For Cartesian Coordinates we define unit vectors lying along the x, y and z axes as $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ respectively and we therefore can write the vector **A** as:

$$\mathbf{A} = A_x\hat{\mathbf{i}} + A_y\hat{\mathbf{j}} + A_z\hat{\mathbf{k}} \quad (1\text{-}9)$$

[We note that this is not the only designation for the base vectors, $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ one will find in the literature. Other common ones are: $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$ and $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$, $\hat{\mathbf{e}}_z$.]

The quantities A_x, A_y and A_z are known as the ***components*** of the vector **A** and they can be found by using the angles α, β and γ shown in Figure 1-3. These angles are known as the ***direction cosines*** of the line segment $|\mathbf{A}|$. In this case the line segment represents the magnitude of the vector. Thus:

$$A_x = A\cos\alpha \tag{1-11}$$

$$A_y = A\cos\beta \tag{1-12}$$

$$A_z = A\cos\gamma \tag{1-13}$$

The magnitude of the vector as can be seen from Figure 1-3 is given by:

$$A = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{1-14}$$

and the direction is determined by the direction cosines.

Because the component approach to representing vectors is usually more useful than representing the vectors through its magnitude and direction we will take this approach in this text almost exclusively. Using components the addition of vectors is defined as follows:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{1-15}$$

$$= (A_x\hat{\mathbf{i}} + A_y\hat{\mathbf{j}} + A_z\hat{\mathbf{k}}) + (B_x\hat{\mathbf{i}} + B_y\hat{\mathbf{j}} + B_z\hat{\mathbf{k}}) \tag{1-16}$$

$$= (A_x + B_x)\hat{\mathbf{i}} + (A_y + B_y)\hat{\mathbf{j}} + (A_z + B_z)\hat{\mathbf{k}} \tag{1-17}$$

$$= C_x\hat{\mathbf{i}} + C_y\hat{\mathbf{j}} + C_z\hat{\mathbf{k}} \tag{1-18}$$

which means that:

$$C_x = A_x + B_x \tag{1-19}$$

$$C_y = A_y + B_y \tag{1-20}$$

$$C_z = A_z + B_z \tag{1-21}$$

Adding vectors becomes equivalent to adding ordinary numbers which means that vectors obey the regular rules of numbers addition. Thus:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \quad \text{(commutative law of addition)} \tag{1-22}$$

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C} \quad \text{(associative law of addition)} \tag{1-23}$$

A negative vector is defined as a vector whose magnitude is the same as the positive vector but whose direction is opposite. In our context vectors of negative magnitudes are not defined. Formally, in components, the negative of the vector $\mathbf{A} = A_x\hat{\mathbf{i}} + A_y\hat{\mathbf{j}} + A_z\hat{\mathbf{k}}$ is $-\mathbf{A} = -A_x\hat{\mathbf{i}} - A_y\hat{\mathbf{j}} - A_z\hat{\mathbf{k}}$.

1-2: Vector Multiplication

Multiplication of the vector **A** by a scalar, c, is simply defined as the multiplication of the magnitude of the vector. Thus, if $\mathbf{B} = c\mathbf{A}$ then:

$$\mathbf{B} = cA_x\hat{\mathbf{i}} + cA_y\hat{\mathbf{j}} + cA_z\hat{\mathbf{k}} \tag{1-24}$$

Multiplication of a vector by a vector is defined to be consistent with the way that physical vector quantities interact. Experimental results indicate that physical vector quantities can combine with one another in two very different ways, one resulting in a scalar and the other in a vector. Vector multiplication which results in a scalar is known as the ***scalar product*** or ***dot product*** and is defined as:

$$\mathbf{A} \bullet \mathbf{B} = AB \cos \psi \tag{1-25}$$

where ψ is the angle between the two vectors. By definition the angle ψ is lies in the range $0 \leq \psi \leq \pi$. We note that the dot product is zero when $\psi = \pi/2$.

The dot product of two vectors, **A** and **B**, can be easily found when they are expressed in component form. Thus since

$$\mathbf{A} = A_x \hat{\mathbf{i}} + A_y \hat{\mathbf{j}} + A_z \hat{\mathbf{k}} \tag{1-26}$$

$$\mathbf{B} = B_x \hat{\mathbf{i}} + B_y \hat{\mathbf{j}} + B_z \hat{\mathbf{k}} \tag{1-27}$$

the scalar product will be:

$$\mathbf{A} \bullet \mathbf{B} = (A_x \hat{\mathbf{i}} + A_y \hat{\mathbf{j}} + A_z \hat{\mathbf{k}}) \bullet (B_x \hat{\mathbf{i}} + B_y \hat{\mathbf{j}} + B_z \hat{\mathbf{k}}) \tag{1-28}$$

$$\mathbf{A} \bullet \mathbf{B} = A_x B_x \hat{\mathbf{i}} \bullet \hat{\mathbf{i}} + A_x B_y \hat{\mathbf{i}} \bullet \hat{\mathbf{j}} + A_x B_z \hat{\mathbf{i}} \bullet \hat{\mathbf{k}} + A_y B_x \hat{\mathbf{j}} \bullet \hat{\mathbf{i}} + A_y B_y \hat{\mathbf{j}} \bullet \hat{\mathbf{j}} + A_y B_z \hat{\mathbf{j}} \bullet \hat{\mathbf{k}} + A_z B_x \hat{\mathbf{k}} \bullet \hat{\mathbf{i}} + A_z B_y \hat{\mathbf{k}} \bullet \hat{\mathbf{j}} + A_z B_z \hat{\mathbf{k}} \bullet \hat{\mathbf{k}} \tag{1-29}$$

The last expression can be greatly simplified. The scalar products $\hat{\mathbf{i}} \bullet \hat{\mathbf{j}}$, $\hat{\mathbf{i}} \bullet \hat{\mathbf{k}}$, $\hat{\mathbf{j}} \bullet \hat{\mathbf{i}}$, $\hat{\mathbf{j}} \bullet \hat{\mathbf{k}}$, $\hat{\mathbf{k}} \bullet \hat{\mathbf{i}}$ and $\hat{\mathbf{k}} \bullet \hat{\mathbf{j}}$ are all equal to zero because the angle between them is $\pi/2$. Similarly $\hat{\mathbf{i}} \bullet \hat{\mathbf{i}}$, $\hat{\mathbf{j}} \bullet \hat{\mathbf{j}}$, and $\hat{\mathbf{k}} \bullet \hat{\mathbf{k}}$ are all equal to one because the magnitude of each vector is one and the angle between them is zero. We therefore have only three non-zero terms and the dot product reduces to:

$$\mathbf{A} \bullet \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \tag{1-30}$$

We can find the magnitude of a vector by using the dot product. From Equation 1-30 we have:

$$\mathbf{A} \bullet \mathbf{A} = A^2 = A_x^2 + A_y^2 + A_z^2 = |\mathbf{A}|^2 \tag{1-31}$$

thus:

$$A = |\mathbf{A}| = \sqrt{\mathbf{A} \bullet \mathbf{A}} \tag{1-32}$$

Finally if we know **A** and **B** we can use the dot product to find the other parameters of the vectors. For example the angle between the vectors will be given by:

$$\cos \psi = \frac{\mathbf{A} \bullet \mathbf{B}}{AB} = \frac{A_x B_x + A_y B_y + A_z B_z}{\sqrt{A_x^2 + A_y^2 + A_z^2}\sqrt{B_x^2 + B_y^2 + B_z^2}} \tag{1-33}$$

In Figure 1-4 we show a geometrical interpretation of the meaning of the dot

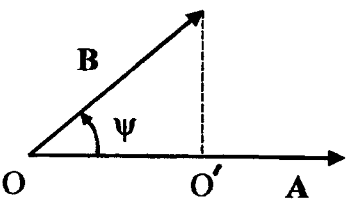


Figure 1-4: Geometric Description of the Scalar Product A•B

product. The distance $\overline{OO'}$ is the projection of **B** upon **A** and it is given by the value of B cos ψ. The projection distance times A is the dot product of the two vectors.

Note that the dot product obeys the commutative and distributive laws of multiplication:

$$\mathbf{A}\bullet\mathbf{B} = \mathbf{B}\bullet\mathbf{A} \tag{1-34}$$

$$\mathbf{A}\bullet(\mathbf{B}+\mathbf{C}) = \mathbf{A}\bullet\mathbf{B} + \mathbf{A}\bullet\mathbf{C} \tag{1-35}$$

but not the associative law, meaning that $\mathbf{A}\bullet(\mathbf{B}\bullet\mathbf{C}) \neq (\mathbf{A}\bullet\mathbf{B})\bullet\mathbf{C}$ because the dot product of a vector with a scalar is not defined.

An important and commonly cited example of a scalar quantity derived from the dot product of two vectors is work. If an object is acted upon by a force, **F**, and is moved a distance, **dr**, then the work done by the force is defined to be $dW = \mathbf{F}(r)\bullet d\mathbf{r}$.

Vector multiplication which results in a vector is known as the ***cross product*** or ***vector product***. It is defined as:

$$\mathbf{C} = \mathbf{A}\times\mathbf{B} = AB\sin\psi\,\hat{\mathbf{n}} \tag{1-36}$$

where again ψ is the angle between the two vectors and $\hat{\mathbf{n}}$ is a unit vector pointing out of the common plane in which the vectors lie. We have illustrated the situation in Figure 1-5.

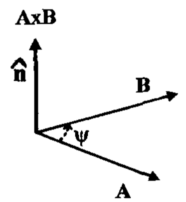


Figure 1-5: Illustration of the Vector Product $\mathbf{C} = \mathbf{A}\times\mathbf{B}$

The direction of $\hat{\mathbf{n}}$ is defined by the right hand rule convention. This convention says that if the fingers of the right hand curve in the same direction as the sense of rotation which would move **A** to **B**, through the smallest angle between them, then the thumb of the right hand points in the direction of positive $\hat{\mathbf{n}}$. We note that:

$$\mathbf{A}\times\mathbf{B} = -\mathbf{B}\times\mathbf{A} \tag{1-37}$$

Again it is convenient to use components to find a general expression for the cross product $\mathbf{A}\times\mathbf{B}$. In terms of our base vectors in Cartesian coordinates, $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ the cross product is:

$$\mathbf{A}\times\mathbf{B} = (A_x\hat{\mathbf{i}} + A_y\hat{\mathbf{j}} + A_z\hat{\mathbf{k}}) \times (B_x\hat{\mathbf{i}} + B_y\hat{\mathbf{j}} + B_z\hat{\mathbf{k}}) \tag{1-38}$$

$$\begin{aligned} &= A_xB_x\hat{\mathbf{i}}\times\hat{\mathbf{i}} + A_xB_y\hat{\mathbf{i}}\times\hat{\mathbf{j}} + A_xB_z\hat{\mathbf{i}}\times\hat{\mathbf{k}} + A_yB_x\hat{\mathbf{j}}\times\hat{\mathbf{i}} + \\ &\quad A_yB_y\hat{\mathbf{j}}\times\hat{\mathbf{j}} + A_yB_z\hat{\mathbf{j}}\times\hat{\mathbf{k}} + A_zB_x\hat{\mathbf{k}}\times\hat{\mathbf{i}} \\ &\quad + A_zB_y\hat{\mathbf{k}}\times\hat{\mathbf{j}} + A_zB_z\hat{\mathbf{k}}\times\hat{\mathbf{k}} \end{aligned} \tag{1-39}$$

This expression can be simplified by using the following relationships:

$$\hat{i}\times\hat{i}=0 \qquad \hat{i}\times\hat{j}=\hat{k} \qquad \hat{i}\times\hat{k}=-\hat{j} \tag{1-40a}$$

$$\hat{j}\times\hat{i}=-\hat{k} \qquad \hat{j}\times\hat{j}=0 \qquad \hat{j}\times\hat{k}=\hat{i} \tag{1-40b}$$

$$\hat{k}\times\hat{i}=\hat{j} \qquad \hat{k}\times\hat{j}=-\hat{i} \qquad \hat{k}\times\hat{k}=0 \tag{1-40c}$$

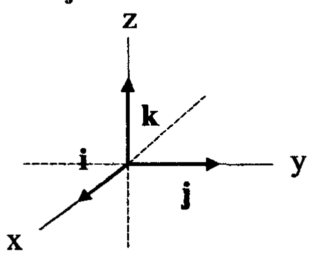


Figure 1-6: The Cross Products of the Unit Vectors

Figure 1-6 can be used as a guide in determining the directions of the cross-products of the Cartesian base vectors. The magnitude is always unity because the base vectors themselves always have magnitude unity. The cross product terms which are zero in Equation 1-40 are due to the fact that $\psi = 0$. The terms which are positive use the fact that $\psi = \pi/2$ and the direction is positive according to the right hand rule. The terms which are negative mean that the direction indicated by the right hand rule is opposite to the sense of the vector. Applying Equation 1-40 to Equation 1-39 gives us:

$$\mathbf{A}\times\mathbf{B} = A_xB_y\hat{k} - A_xB_z\hat{j} - A_yB_x\hat{k} + A_yB_z\hat{i} + A_zB_x\hat{j} - A_xB_y\,\hat{k} \tag{1-41}$$

$$= (A_yB_z - A_zB_y)\hat{i} + (A_zB_x - A_xB_z)\,\hat{j} + (A_xB_y - A_yB_x)\,\hat{k} \tag{1-42}$$

which can be represented by the determinant:

$$\mathbf{A}\times\mathbf{B} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \tag{1-43}$$

A geometrical interpretation of the cross product is that it represents the area of the parallelogram enclosed by the vectors **A** and **B**. To see this we note from Figure 1-7 that the area of this parallelogram is given by the product of the base, in this case $|\mathbf{A}|$, times the

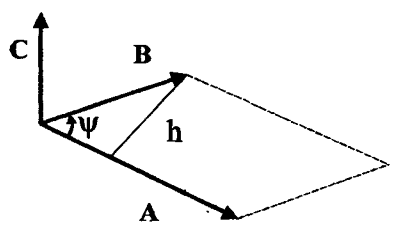


Figure 1-7: Geometric Definition of the Cross Product

height, h, which is $|\mathbf{B}| \sin \psi$. Thus the area is AB $\sin \psi$ which is also the magnitude of the vector **C**. The area vector **S** which is, in this case, equivalent to the vector product **C** has its direction defined by the right hand rule, meaning that it is in the direction the thumb would point if the fingers are curling from **A** to **B**.

It should be noted that the vector product is not a "true" vector. For example, vectors under coordinate rotations maintain their properties, meaning that their magnitudes and directions do not change. Under some types of transformations, however, they do not maintain these properties. For example, if we make an inversion transformation, in other words transforming $A_x \rightarrow -A_x$, $B_x \rightarrow -B_x$, and so on, then our new vectors will be $\mathbf{A}' = -\mathbf{A}$ and $\mathbf{B}' = -\mathbf{B}$. We see, however, that the vector, $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, does not transform under inversion like a vector, because:

$$\mathbf{C}' = \mathbf{A}' \times \mathbf{B}' \tag{1-44}$$

$$= (-\mathbf{A}) \times (-\mathbf{B}) \tag{1-45}$$

$$= \mathbf{A} \times \mathbf{B} = \mathbf{C} \tag{1-46}$$

If it were a true vector then $\mathbf{C}'$ would be equal to $-\mathbf{C}$. Since it has different transformation properties than a regular vector, it is called a ***pseudovector*** or ***axial vector***. Technically it is a second rank tensor, a mathematical quantity we will discuss later in this chapter.

Other combinations of vector multiplication frequently occur. One possible combination is called the ***scalar triple product***, which can be represented by a determinant as:

$$\mathbf{A}\bullet(\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix} \tag{1-47}$$

We have noted that $\mathbf{A} \times \mathbf{B}$ is a pseudovector because it does not change sign when its coordinates are inverted. We note also that Equation 1-47 defines a "scalar" $\mathbf{A}\bullet(\mathbf{B} \times \mathbf{C})$ which does change sign upon inversion of the coordinates. Thus it does not act like a true scalar. It is called a ***pseudoscalar***, also another form of a second rank tensor.

Another possible combination is known as the ***triple cross product***, which can be represented as:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}\,(\mathbf{A}\bullet\mathbf{C}) - \mathbf{C}\,(\mathbf{A}\bullet\mathbf{B}) \tag{1-48}$$

As a mnemonic device the above expression is often called the "BAC minus CAB" rule.

Note that $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \neq (\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$, and thus the cross product, does not obey the associative law for multiplication.

As in the case of the scalar product the definition of the vector product comes about because it describes mathematically what we determine through experiment.

One well known situation where vector quantities combine according to the vector product occurs in the case of a rotating wheel as illustrated in Figure 1-8.

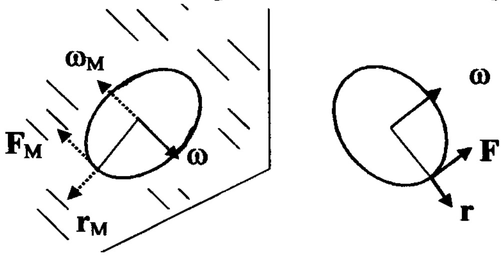


Figure 1-8: Mirror Image Transformation of the Angular Velocity Vector

By convention ω, the angular velocity, is a vector which points along the axis of rotation in a direction given by the right hand rule, meaning that if the fingers point along the direction of rotation of the rim then the thumb will point along the angular velocity vector. If a force, **F**, acts at a point on the perimeter of the wheel then it is found that the angular velocity will increase or decrease according to the equation:

$$\mathbf{r} \times \mathbf{F} = \mathrm{I}\, d\omega/DT \qquad (1\text{-}49)$$

where I is the moment of inertia of the wheel. Thus the vector ω will change its magnitude due to a change in vectors operating at right angles to it. We note in Figure 1-8 that the mirror images of the vectors obey the expected rules for reflection. However, we see that the mirror image of ω, denoted as ω_M in Figure 1-8, is not along the direction given by the cross-product of $\mathbf{r}_M \times \mathbf{F}_M$ which instead is the vector ω' in Figure 1-8. This is a good geometric description of how a pseudovector's behavior in reflection differs from a regular vector.

One might naturally continue our discussion of vectors with the division of vectors. However, except for division by a scalar, we will not define vector division, since it has no application in the topics we will cover.

Problem 1-1

Given the vectors:

$$\mathbf{A} = \hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}} \quad \mathbf{B} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} - \hat{\mathbf{k}} \quad \mathbf{C} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 2\hat{\mathbf{k}}$$

find:

(a) $\mathbf{A}+\mathbf{B}$ (b) $\mathbf{B}+\mathbf{A}$ (c) $\mathbf{A}+\mathbf{B}+\mathbf{C}$ (d) $\mathbf{A}\bullet\mathbf{B}$ (e) $\mathbf{A}\bullet\mathbf{C}$

(f) $\mathbf{B}\bullet\mathbf{C}$ (g) $\mathbf{A}\times\mathbf{B}$ (h) $\mathbf{B}\times\mathbf{A}$ (i) $\mathbf{A}\times\mathbf{C}$ (j) $\mathbf{B}\times\mathbf{C}$

(k) $\mathbf{A}\bullet(\mathbf{B}\times\mathbf{C})$ (l) $\mathbf{B}\bullet(\mathbf{A}\times\mathbf{C})$ (m) $\mathbf{C}\bullet(\mathbf{A}\times\mathbf{B})$ (n) $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$

(o) $(\mathbf{A}\times\mathbf{B})\times\mathbf{C}$ (p) $|\mathbf{A}|$ (q) $|\mathbf{B}|$ (r) $|\mathbf{C}|$

(s) $\cos\alpha$ where $\alpha = \angle(\mathbf{A},\mathbf{B})$ (t) $\cos\beta$ $[\beta = \angle(\mathbf{B},\mathbf{C})]$

(u) $\cos\gamma$ where $\gamma = \angle(\mathbf{A},\mathbf{C})$ (v) $(\mathbf{A}\times\mathbf{B})\bullet(\mathbf{A}\times\mathbf{C})$

Problem 1-2
Show that $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A}\bullet\mathbf{C}) - \mathbf{C}(\mathbf{A}\bullet\mathbf{B})$ by expanding the triple vector product in Cartesian Coordinates.

1-3: Fields

Most generally, physical quantities are not restricted to having values just at specific points. Instead the most usual and useful situation is when the description of the physical quantity involves a complete mathematical description over all space. This description may be functional or simply numerical. Such a representation is called a ***field***.

Just as there are scalar and vector quantities there are scalar and vector fields which represent them. In Figure 1-9 we show how fields can be described through

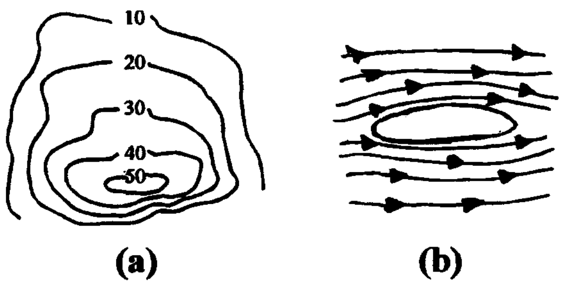


Figure 1-9: Diagrams of: (a) A Scalar Field and (b) A Vector Field

diagrams. In Figure 1-9(a) we use a topographical map as our example of a scalar field. On such a map the lines indicate points of equal elevation. Each point on the line has the same value and that value completely describes the quantity we are interested in, namely the elevation at any point on the map.

In Figure 1-9(b) we have used as our example the flow of fluid over an air foil. Now the lines mean something entirely different. Each line, called a streamline, represents a set of points where a vector will lie tangent to the line connecting the points and indicates the direction the fluid is flowing. The magnitude of the vector is represented by the closeness of the lines. The closer they are together the higher the strength of the vector field as we have noted in the figure.

The field concept has shown to be a most successful way to describe and predict experimental physical phenomena. Our aim is almost always to find a mathematical function which describes any given physical situation and represents the impact of the physical entity over all space. When we do that we can plot the field, describe how it interacts with other fields or objects or interpret the source or sources of the field. If our field is scalar we would describe it in Cartesian coordinates as $U = U(x,y,z)$. On the other hand if we have a vector field its description in

Cartesian coordinates would be $\mathbf{A}(x,y,z) = A_x(x,y,z)\hat{\mathbf{i}} + A_y(x,y,z)\hat{\mathbf{j}} + A_z(x,y,z)\hat{\mathbf{k}}$. A complete description of the gravitational field of the earth would enable us at any point to find the magnitude and direction of the force on an object from which we could determine its velocity and trajectory, a technique used to predict the motion of satellites, for example. This makes it unnecessary to calculate continually the force at each point using the distance between the earth and the object. Because the field concept with respect to the electric and magnetic field has led to substantial simplification we will use the field concept as we develop the theory of electricity and magnetism. In fact the action-at-a distance concept has no use when we investigate electromagnetic waves.

1-4: Vector Calculus: The Gradient

Our objective is to use the physical situation we are given to find the field which represents the physical phenomena demonstrated by our experiment. Thus at any given point in space we will have a complete mathematical description of the effect be it scalar or vector. Often these fields can be described by continuous functions whose derivatives or integrals are useful in obtaining new information. Therefore we often will apply the techniques of calculus to a given field representation.

As a first example we will introduce the concept of the ***gradient***. If we take a scalar function, U(x,y,z), there will be many occasions when we will want to know its directional derivative, dU/dl, where dl is the incremental distance along any line of arbitrary orientation. We know from calculus that the total derivative can be written as:

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz \tag{1-50}$$

From the form of the scalar product (Equation 1-30) we can see that Equation 1-50 can also be written as the scalar product of two vectors. One, which we will call ∇U, given by:

$$\nabla U = \frac{\partial U}{\partial x}\hat{\mathbf{i}} + \frac{\partial U}{\partial y}\hat{\mathbf{j}} + \frac{\partial U}{\partial z}\hat{\mathbf{k}} \tag{1-51}$$

and the other, the distance increment, $d\mathbf{l}$, given by:

$$d\mathbf{l} = dx\,\hat{\mathbf{i}} + dy\,\hat{\mathbf{j}} + dz\,\hat{\mathbf{k}} \tag{1-52}$$

We note that Equation 1-52 represents the increment of distance between a point on the surface $U(x,y,z) = \Phi_1$ and a point on the surface $U(x + dx, y + dy, z + dz) = \Phi_2$. Thus:

$$dU = \nabla U \bullet d\mathbf{l} = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz \tag{1-53}$$

The vector described by Equation 1-51 can be represented as a combination of a vector operator and the scalar function U(x, y, z). By vector operator we mean in this specific case that it combines with other functions by differentiating them. It is not a multiplicative factor. We call ∇ the ***del operator*** and it is defined as:

$$\nabla = \frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}} \tag{1-54}$$

and the expression, ∇U, is called the gradient of the scalar field U.

We will now show that the gradient has three important properties which can be derived from the equation:

$$dU = \nabla U \cdot dl = |\nabla U||dl|\cos\alpha \qquad (1\text{-}55)$$

We have plotted in Figure 1-10 a scalar field described by the equation $U(x,y,z) = \Phi$ for two different values of Φ. For the range of values of x, y, z there will be a series of sheets of which in Figure 1-10 we illustrate only two.

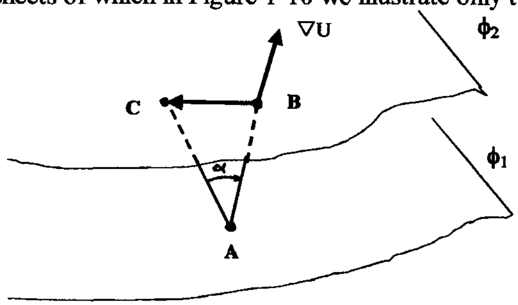


Figure 1-10: Behavior of ∇ U(x,y,z)

The incremental change in U, moving from one sheet to the other, is:

$$dU = U_1(x,y,z) - U_2(x,y,z) \qquad (1\text{-}56)$$

We can see from Figure 1-10 that there would be an infinite number of paths going from the point A on $U_1(x,y,z)$ to the adjacent surface $U_2(x,y,z)$.

We can show that:

(a) $|\nabla U|$ is equal to the maximum value of the directional derivative dU/dl.

This follows directly from Equation 1-55 above. Rewriting our equation (since $|dl| = dl$), we have:

$$\frac{dU}{dl} = |\nabla U|\cos\alpha \qquad (1\text{-}57)$$

The maximum value of dU/dl occurs when $\cos\alpha = 1$, which simply means that for this case $dU/dl = |\nabla U|$ and therefore $|\nabla U|$ is equal to the maximum rate of change of the directional derivative dU/dl.

(b) The direction of ∇U is normal to the surface defined by the function, $U(x,y,z) = \Phi$.

To show this we take two points, B and C, on the sheet $U_2(x,y,z)$ as noted in Figure 1-10. Because we are confined to that sheet, $dU = 0$ and thus Equation 1-55 becomes:

$$\nabla U \cdot dl_{BC} = |\nabla U||dl_{BC}|\cos\alpha = 0 \qquad (1\text{-}58)$$

Again since ∇U and dl_{BC} are not in general zero at arbitrary points on the sheet this equation can only be true if $\cos\alpha = 0$, which only happens when $\alpha = \pi/2$. Thus ∇U must be perpendicular to U since dl_{BC} lies on the surface defined by $U_2(x,y,z)$.

From this result we can find a vector normal to any surface, represented by the equation $U(x,y,z) - c = 0$. Since ∇U is normal to the surface then the normal vector to that surface, usually defined as $\hat{\mathbf{n}}$, will be given by the expression:

$$\hat{\mathbf{n}} = \frac{\nabla U}{|\nabla U|} \tag{1-59}$$

For the special case where dl lies along $\hat{\mathbf{n}}$, we can designate dl as the **normal derivative**, dn. From Equation 1-55:

$$dU = \nabla U \cdot d\mathbf{l} = \nabla U \cdot d\mathbf{n} = \nabla U \cdot \hat{\mathbf{n}}\, dn \tag{1-60}$$

Thus:

$$\frac{dU}{dn} = \nabla U \cdot \hat{\mathbf{n}} = \nabla U \cdot \frac{\nabla U}{|\nabla U|} = |\nabla U| \tag{1-61}$$

which means that the magnitude of $|\nabla U|$ is equal to the normal derivative of U.

(c) The direction of ∇U lies in the direction of the maximum rate of increase of U.

We have shown that ∇U is along the direction of $\hat{\mathbf{n}}$. We have also shown that $|\nabla U|$ is the maximum value of dU/dl at any point on U. From Equation 1-57 this means that $\cos \alpha$ is unity and its maximum value. Because the sign is positive ∇U must then be pointing in the direction of increasing U and be the maximum rate of increase.

The gradient can be visualized in another way. Suppose you are climbing up the side of a hill and want to follow the gradient of the hill. If you are walking up the steepest direction at each point in your path up the hill you would be following the gradient of the surface of the hill.

The gradient has another useful property. Suppose we integrate both sides of Equation 1-60 giving us:

$$\int_1^2 \nabla U \cdot d\mathbf{l} = \int_1^2 dU = U_2 - U_1 \tag{1-62}$$

On the right side of the equation we have what is known as a total integral whose value can be found simply by evaluating U at the end points. However on the left side we have a special type of integral known as a line integral. In general, given specific limits, this kind of integral will vary with the path of integration. However Equation 1-62 says that the value of the integral is independent of the path which means that in this restricted case the line integral is the same as the regular integral. Since this is not generally true it is worthwhile to investigate the meaning of Equation 1-62 more thoroughly. This will tell us more about the properties of the gradient.

The line integral can be converted to a regular set of integrals by using components. For example, in Cartesian coordinates we have:

$$\int_l \mathbf{A} \cdot d\mathbf{l} = \int_l A\,dl \cos \alpha = \int A_x(x,y,z)\, dx + \int A_y(x,y,z)\, dy + \int A_z(x,y,z)\, dz \tag{1-63}$$

Here α is the angle between the direction of the vector field **A** and the line increment d**l**.

We have illustrated in Figure 1-11 how the direction of the line and the vector field will vary over the path of integration. We can see that in general the form of the line integral noted in Equation 1-63 is quite different than the scalar integral which would merely be $\int A\, dl$.

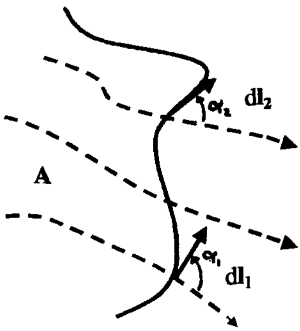


Figure 1-11: The Integrand A•dl Changes Along the Path of Integration

Example 1-1 Integration of a Line Integral:Independence of Path

As a simple example of a line integral we integrate the gradient of the function

$$U = xy \qquad (1\text{-}64)$$

along a series of paths. The gradient is:

$$\nabla U = y\hat{i} + x\hat{j} \qquad (1\text{-}65)$$

We will use the three different paths illustrated in Figure 1-12. Path (1)

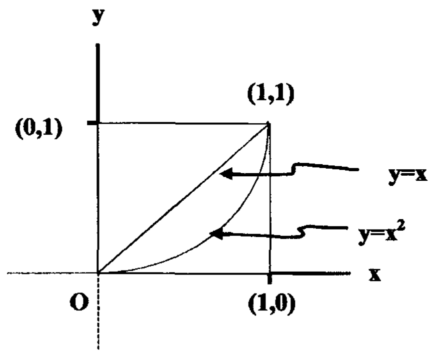


Figure 1-12: Paths of Integration of a the Line Integral

will be the straight paths (0,0) → (1,0) → (1,1). Path (2) will be the path y = x, and finally path (3) will be the parabola $y = x^2$. All go through the end points (0,0) and (1,1). Over path (1) our integral will be:

$$\int_{0,0}^{1,1} \nabla U \bullet dl = \int_{0,0}^{1,0} y\, dx + \int_{1,0}^{1,1} x\, dy = 0 + \int_0^1 dy = 1 \qquad (1\text{-}66)$$

where we have used the fact that y = 0 on the line segment (0,0) → (1,0) and x = 1 on the line segment (1,0) → (1,1). If we next go on to path (2) we have the integral:

$$\int_0^1 x\,dx + \int_0^1 y\,dy = \left[\ \frac{x^2}{2}\ \right]_0^1 + \left[\ \frac{y^2}{2}\ \right]_0^1 = \frac{1}{2} + \frac{1}{2} = 1 \qquad (1\text{-}67)$$

Next we can use path (3) . Our integral then becomes:

$$\int_0^1 x^2 dx + \int_0^1 y^{1/2}\,dy = \left[\ \frac{x^3}{3}\ \right]_0^1 + \left[\ \frac{2}{3} y^{3/2}\right]_0^1 + \frac{1}{3} + \frac{2}{3} = 1 \qquad (1\text{-}68)$$

thus our integral is independent of the path chosen. (Note that this does not prove that the value of the integral is independent of all possible paths which can be drawn between the limits. However the fact that it works for these paths is a good indication that indeed our integral is independent of path.) This property means that if our integration takes us around a complete loop then our integral would be:

$$\oint_C \nabla U \cdot \mathbf{dl} = 0 \qquad (1\text{-}69)$$

This can be seen directly from our above integrals. On path (1) our value of the integral was 1. On path (2) it was also 1. If we were to do a complete cycle our sense of integration would be in the opposite direction along $y = x$ and our value for the integral would be -1 making the total value of the integral around the complete circuit zero.

Now suppose we find the line integral of the vector function, $\mathbf{F} = y^2\hat{\mathbf{i}} + x\hat{\mathbf{j}}$. Now our integral will be:

$$\int_{0,0}^{1,1} \mathbf{F}\cdot\mathbf{dl} = \int_{0,0}^{1,0} y^2 dx + \int_{1,0}^{1,1} x\,dy = 0 + \int_0^1 dy = 1 \qquad (1\text{-}70)$$

For the line $y = x$, our integral is:

$$\int_{0,0}^{1,1} \mathbf{F}\cdot\mathbf{dl} = \int_0^1 x^2 dx + \int_0^1 y\,dy = \left[\ \frac{x^3}{3}\ \right]_0^1 + \left[\ \frac{y^2}{2}\ \right]_0^1 = \frac{1}{3} + \frac{1}{2} = \frac{5}{6} \qquad (1\text{-}71)$$

Next we will use the line $y = x^2$:

$$\int_{0,0}^{1,1} \mathbf{F}\cdot\mathbf{dl} = \int_0^1 x^4 dx + \int_0^1 y^{1/2} dy = \left[\frac{x^5}{5}\ \right]_0^1 + \left[\ \frac{2}{3} y^{3/2}\right]_0^1 = \frac{1}{5} + \frac{2}{3} = \frac{13}{15} \qquad (1\text{-}72)$$

Thus this line integral is not independent of path and

$$\oint \mathbf{F}\cdot\mathbf{dl} \neq 0 \qquad (1\text{-}73)$$

The line integral of a vector function, **F**, will be independent of path if **F** is the gradient of a scalar function ϕ, in other words, whether or not $\mathbf{F} = \nabla\phi$. In this case then the integral is:

$$\int_A^B \mathbf{F}\cdot\mathbf{dl} = \int_A^B \nabla\phi\cdot\mathbf{dl} = \int_A^B d\phi = \phi(B) - \phi(A) \qquad (1\text{-}74)$$

In principle all we need to do is determine if the integrand is the gradient of some function and then evaluate the function at the end points. In reality, although there are many techniques to do this, it may be difficult to find the function. Note that there is no function, f, for which $\mathbf{F} = y^2\,\hat{\mathbf{i}} + x\hat{\mathbf{j}}$ is the gradient. We will see that the property of being independent of path has important implications.

More generally we can say that the integral $\int_A^B F\,dx + G\,dy$ is independent of path if there is a function P such that $\partial P/\partial x = F(x,y)$ and $\partial P/\partial y = G(x.y)$.

1-5: The Divergence of a Vector

We have found the result of what happens when the del operator operates on a scalar field. Now we consider what happens when we take a del operation on a vector field. We have two possibilities; a dot product operation, and a cross product operation. Both are defined and find extensive use.

We first discuss the dot product. If we have a vector field described by $\mathbf{A}(x,y,z) = A_x(x,y,z)\hat{\mathbf{i}} + A_y(x,y,z)\hat{\mathbf{j}} + A_z(x,y,z)\hat{\mathbf{k}}$ then we define the ***divergence*** of the vector field as:

$$\nabla\cdot\mathbf{A} = \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right)\cdot\left(A_x\hat{\mathbf{i}} + A_y\hat{\mathbf{j}} + A_z\hat{\mathbf{k}}\right)$$

$$= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \qquad (1\text{-}75)$$

Although this is a straightforward mathematical definition, to describe its physical meaning we introduce another physical quantity which is known as the ***flux***. Given a vector quantity **A**, the flux, Ψ_A, of the vector is defined to be:

$$\Psi_A = \int \mathbf{A}\cdot d\mathbf{S} \qquad (1\text{-}76)$$

where **dS** is the area increment vector through which the vector **A** "flows". Thus flux is the measure of the number of field vector lines which cut through per unit surface area.

In order to carry out a surface integral we need to define a means to determine the direction of the area vector in a consistent manner. In addition it has to coordinate with the way we define the positive direction of the line increment on its boundary. [This becomes important when we discuss Stokes's Theorem in Section 1-8 below.] If the surface is closed the traversal along the boundary is usually defined so that the area vector points in the direction of the outward normal of the surface away from inside of the surface. For a plane surface, or a surface which has no definable interior, this definition would make the direction of the area vector ambiguous. Therefore we relate the area vector to the positive direction of the line increment, **dl**, according to the right hand rule.

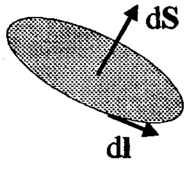


Figure 1-13: Definition of the Direction of the Area Vector

In Figure 1-13 the directions of both the area vector and **dl** are indicated. If the fingers of the right hand curl along the direction **dl** then the thumb points in the direction of the area vector which represents that area lying inside the boundary. Otherwise the definition of the area vector is quite arbitrary and we can set the direction in a way which is most convenient for the parameters of the problem.

To see how flux relates to the divergence we select an example from fluid mechanics. Suppose we have a fluid of density, ρ, moving with velocity **v** through a surface of area dS, as indicated in Figure 1-14. If the fluid flows a distance dy in

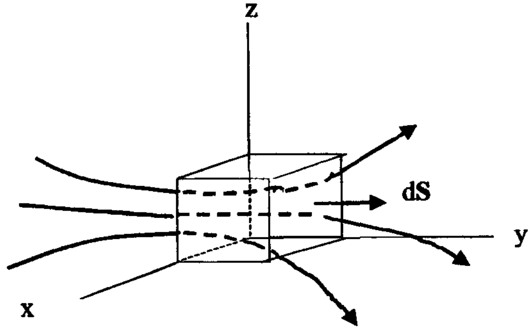


Figure 1-14: Fluid Flowing Through a Surface

time dt, in terms of the general velocity, **v**, we have:

$$dy = \mathbf{v} \cdot \hat{\mathbf{j}}\, dt \tag{1-77}$$

Then the incremental amount of fluid flowing past the surface is:

$$dm = \rho\, dx\, dy\, dz = \rho\, dx(\mathbf{v} \cdot \hat{\mathbf{j}}\, dt)\, dz = \rho\mathbf{v} \cdot (dx\, dz\, \hat{\mathbf{j}})\, dt = \rho\mathbf{v} \cdot d\mathbf{S}\, dt \tag{1-78}$$

Thus the total amount of fluid flowing through the surface each second is just the integral:

$$\frac{dM}{dt} = \int_S \rho\mathbf{v} \cdot d\mathbf{S} \tag{1-79}$$

which very much looks like our flux equation. In this case the flux represents the rate of fluid flow:

$$\Psi_F = \int_S \rho\mathbf{v} \cdot d\mathbf{S} \tag{1-80}$$

Here we identify the flux, Ψ_F, as the flux of the vector field, $\rho\mathbf{v}$ which by Equation 1-79 also is equivalent to the movement of mass per unit time which flows through the surface.

1-6: The Divergence Theorem

We next show that the flux of a vector field is related to its divergence. Let us take the example of fluid flow again and generalize to the case where we have fluid flowing through the surface of a volume in an arbitrary direction.

As before the measure of fluid flow is:

$$\mathbf{A}(x,y,z) = \rho\mathbf{v} \tag{1-81}$$

Let us divide the region in which the fluid is flowing into an increment of volume of dimension $\Delta x \Delta y \Delta z$ as we have illustrated in Figure 1-15. We will determine the flux through the surface of the cube. At the surface designated as (1) in the diagram the coordinates are x, y and z and the flux is:

$$\mathbf{A} \cdot \Delta\mathbf{S} = -A_y(x,y,z)\Delta x \Delta z \tag{1-82}$$

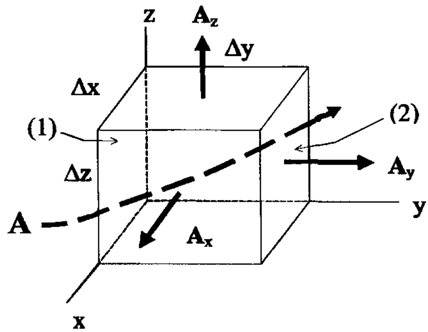


Figure 1-15: Flux Flowing through a Cube

At the surface designated as (2) in the diagram the coordinates are x, y + Δy, z and the flux is:

$$\mathbf{A}\bullet\Delta\mathbf{S} = A_y(x,y+\Delta y,z)\Delta x\Delta z \tag{1-83}$$

Thus the change of flux through the surface is:

$$\Delta\Psi_A)_y = [A_y(x,y+\Delta y,z) - A_y(x,y,z)]\ \Delta x\Delta z \tag{1-84}$$

$$\Delta\Psi_A)_y = \frac{[A_y(x,y+\Delta y,z)-A_y(x,y,z)]}{\Delta y}\Delta x\Delta y\Delta z \tag{1-85}$$

which, as Δx, Δy, Δz → 0, becomes:

$$d\Psi_A\Big)_y = \frac{\partial A_y}{\partial y}\ dx\ dy\ dz \tag{1-86}$$

Similarly, through the other sides we will have, in the same way:

$$d\Psi_A\Big)_x = \frac{\partial A_x}{\partial x}\ dx\ dy\ dz \tag{1-87}$$

$$d\Psi_A\Big)_z = \frac{\partial A_z}{\partial z}\ dx\ dy\ dz \tag{1-88}$$

and the total change in flux is given by:

$$d\Psi_A = \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right)\ dx\ dy\ dz \tag{1-89}$$

$$= (\nabla\bullet\mathbf{A})\ d\tau \tag{1-90}$$

but since we have defined flux as $d\Psi_A = \mathbf{A}\bullet d\mathbf{S}$, we have the equation:

$$\Psi_A\ =\ \oint_S \mathbf{A}\bullet d\mathbf{S} = \int_\tau \nabla\bullet\mathbf{A}\ d\tau \tag{1-91}$$

where the surface integral includes the entire closed surface. Equation 1-91 is known as the ***divergence theorem***.

To see what this means for a physical situation we will use our example of fluid flow once more. Our "flow vector" was shown to be $\mathbf{A} = \rho\mathbf{v}$. If the rate of fluid flowing into the surface is the same as the rate flowing out of the surface then the flux integral $\oint\mathbf{A}\bullet d\mathbf{S}$ must be zero. Since we are talking about an arbitrary surface this also means that $\nabla\bullet\mathbf{A}$ is zero. If there is more fluid flowing out of the surface than flowing into the surface then $\oint\mathbf{A}\bullet d\mathbf{S}\ \neq\ 0$ and therefore $\nabla\bullet\mathbf{A}\ \neq\ 0$. Conservation of matter leads us to assume that the only way this situation could occur is that there must be creation of fluid within the enclosed volume. For any vector field, therefore, a non-zero divergence means that there is a source of the vector field somewhere.

Here we have emphasized the physical meaning of Equation 1-91. However, the divergence theorem is also useful in applied mathematics because it is a way to evaluate either surface integrals or volume integrals. Usually the surface integral is the more difficult because it is a two-dimensional scalar product. In Figure 1-16

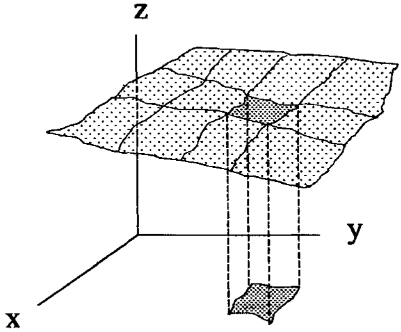


Figure 1-16: Evaluation of a Surface Integral can be Done by Converting the Integral over a Surface to a Two-dimensional Integral over the x-y Plane

we indicate geometrically how a surface integral may be evaluated. An integral over the surface would be

$$\int \mathbf{A}\cdot d\mathbf{S} = \int A\, dS \cos\alpha = \int \mathbf{A}\cdot\hat{\mathbf{n}}\, dS \tag{1-92}$$

As we noted in Section 1-4 we know that the normal to the surface is along the gradient. If the surface is written in the functional form $U(x,y,z) - c = 0$ then the normal vector to the surface at any point can be written as:

$$\hat{\mathbf{n}} = \frac{\nabla U}{|\nabla U|} \tag{1-93}$$

and our surface integral becomes:

$$\int \mathbf{A}\cdot d\mathbf{S} = \int \mathbf{A}\cdot \frac{\nabla U}{|\nabla U|}\, dS \tag{1-94}$$

if we take the projection of the surface onto the x-y plane, as we have illustrated in Figure 1-16, we would have:

$$dx\, dy = \hat{\mathbf{k}}\cdot\hat{\mathbf{n}}\, dS = \hat{\mathbf{k}}\cdot \frac{\nabla U}{|\nabla U|}\, dS = \hat{\mathbf{k}}\cdot \frac{\frac{\partial U}{\partial x}\hat{\mathbf{i}} + \frac{\partial U}{\partial y}\hat{\mathbf{j}} + \frac{\partial U}{\partial z}\hat{\mathbf{k}}}{|\nabla U|}\, dS = \frac{\left|\frac{\partial U}{\partial z}\right|}{|\nabla U|}\, dS \tag{1-95}$$

thus:

$$dS = \frac{|\nabla U|}{|\partial U/\partial z|}\, dx\, dy \tag{1-96}$$

and the final form of the integral is

$$\int \mathbf{A}\cdot d\mathbf{S} = \int \frac{\mathbf{A}\cdot\nabla U}{|\partial U/\partial z|}\, dx\, dy \tag{1-97}$$

Thus we have converted our surface integral into an integral in the x-y plane. If it is more convenient to use projections upon other planes the equivalent expressions would be:

$$\int \mathbf{A} \bullet d\mathbf{S} = \int \frac{\mathbf{A} \bullet \nabla U}{|\partial U/\partial x|} \, dy \, dz = \int \frac{\mathbf{A} \bullet \nabla U}{|\partial U/\partial y|} \, dx \, dz \tag{1-98}$$

In some situations this process makes finding of the surface integral easier. However, in many cases the integrals represented by Equations 1-97 and 1-98 still will not reduce to common integrals and nothing is gained in the transformation.

Example 1-2: Surface Integration

In order to show how a surface integral can be evaluated we will take a relatively simple situation and take the flux of a vector field, $\mathbf{A} = x\hat{i}$, intersecting with a surface defined by the plane $U = x + y + z - 1 = 0$ and the x-y, x-z and y-z planes such that $x \leq 0, y \leq 0, z \leq 0$, as shown in Figure 1-17.

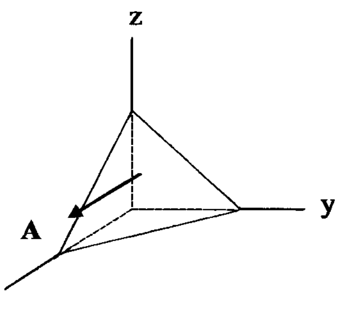


Figure 1-17: The Vector Field, $\mathbf{A} = x\hat{i}$, Intersects the Plane $U = x + y + z - 1$

The parameters of the equation are:

$$\nabla U = \hat{i} + \hat{j} + \hat{k} \tag{1-99}$$

$$\mathbf{A} \bullet \nabla U = x \tag{1-100}$$

$$\nabla \bullet \mathbf{A} = 1 \tag{1-101}$$

We will project the sloping surface onto the x-y plane. Thus we use:

$$|\partial U/\partial z| = 1 \tag{1-102}$$

Getting the equation:

$$\int \mathbf{A} \bullet d\mathbf{S} = \int \frac{\mathbf{A} \bullet \nabla U}{|\partial U/\partial z|} \, dx \, dy = \int x \, dx \, dy \tag{1-103}$$

Thus:

$$\int \frac{\mathbf{A} \bullet \nabla U}{|\partial U/\partial z|} \, dx \, dy = \int_0^1 x \, dx \int_0^{1-x} dy = \int_0^1 (x - x^2) \, dx = \frac{1}{6} \tag{1-104}$$

The value of the surface integrals on the other three surfaces is zero. For the x-y plane and the x-z plane the integrals are zero because the terms **A** and **dS** are perpendicular to one another on those two planes. The value of the integral on the y-z plane is zero because x = 0 and thus **A** = 0 on that surface. We must include the entire surface in our surface integration, which we have done.

We use the divergence theorem to check our answer. Thus:

$$\int_\tau \nabla\cdot\mathbf{A}\, d\tau = \int dx\, dy\, dz = \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} dz \tag{1-105}$$

$$\int_\tau \nabla\cdot\mathbf{A}\, d\tau = \int_0^1 \tfrac{1}{2}(1 - 2x + x^2)\, dx = \left[\tfrac{1}{2}\, x - \tfrac{x^2}{2} + \tfrac{x^3}{6}\right]_0^1 = \tfrac{1}{6} \tag{1-106}$$

and our answers are consistent, showing for this specific example, at least, that the integral $\int\nabla\cdot\mathbf{A}d\tau$ is equal to the surface integral, $\oint\mathbf{A}\cdot d\mathbf{S}$. The fact that $\nabla\cdot\mathbf{A} \neq 0$ means that the field has a source somewhere within the volume.

1-7: The Curl of a Vector

We next consider how the del operator combines with a vector using the cross product. Using the determinant representation of a vector product (Equation 1-43) we have:

$$\nabla \times \mathbf{A} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \tag{1-107}$$

$$= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\hat{\mathbf{k}} \tag{1-108}$$

We call this function the ***curl*** of **A**.

To get some idea of what this quantity means in a physical situation we examine the two vector fields:

$$\mathbf{A}_1 = -y\hat{\mathbf{i}} + x\hat{\mathbf{j}} \tag{1-109}$$

and

$$\mathbf{A}_2 = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} \tag{1-110}$$

which we plot in Figure 1-18.

From the definition of the curl and the divergence, operating upon **A**, we have:

$$\nabla \times \mathbf{A}_1 = \left(\frac{\partial A_{1z}}{\partial y} - \frac{\partial A_{1y}}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial A_{1x}}{\partial z} - \frac{\partial A_{1z}}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial A_{1y}}{\partial x} - \frac{\partial A_{1x}}{\partial y}\right)\hat{\mathbf{k}} \tag{1-111}$$

$$= (0 - 0)\hat{\mathbf{i}} + (0 - 0)\hat{\mathbf{j}} + (1 + 1)\hat{\mathbf{k}} \tag{1-112}$$

$$\nabla \times \mathbf{A}_1 = 2\hat{\mathbf{k}} \tag{1-113}$$

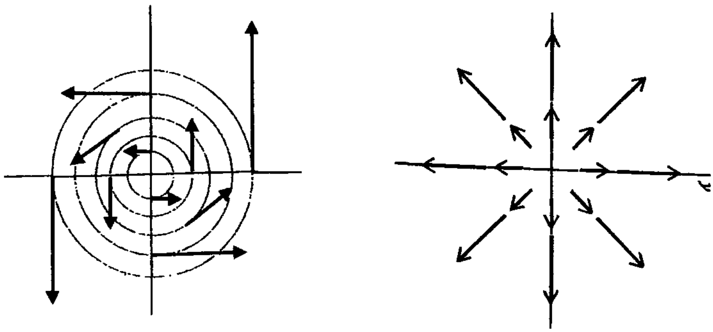
$$\nabla\cdot\mathbf{A}_1 = \frac{\partial A_{1x}}{\partial x} + \frac{\partial A_{1y}}{\partial y} + \frac{\partial A_{1z}}{\partial z} = 0 \tag{1-114}$$

Using these same operations on $\mathbf{A}_2$, we have:

$$\nabla \times \mathbf{A}_2 = 0 \tag{1-115}$$

and,

$$\nabla \cdot \mathbf{A}_2 = 2 \tag{1-116}$$



(a) (b)

Figure 1-18 Plot of: (a) $A_1 = -y\hat{i} + x\hat{j}$ and (b) $A_2 = x\hat{i} + y\hat{j}$

Thus $\mathbf{A}_1$ has a curl and no divergence and $\mathbf{A}_2$ has a divergence but no curl. From Figure 1-18(a), the plot of $\mathbf{A}_1$, it is evident that what this means physically is that the curl is related to the circulation of the vector field about a point. Using the analogy of fluid flow again it is a measure of the vortices which may occur in the field. Indeed the name curl is a description of what is happening to the field meaning that the field lines curve. In some texts, especially European ones, $\nabla \times \mathbf{A}$ is called rot **A** also a descriptive term for what the term describes physically.

From Figure 1-18(b) since there is no vortex motion one would expect the curl of the field to be zero and it is. The divergence of the field is not. Again the term is very descriptive. The field diverges as we move away from the origin, and the nature of the field indicates that there must be a source somewhere in order to create the flow as diagrammed. This situation would be what would occur if a point source of fluid is flowing out in all directions from the origin onto an infinite smooth plane.

1-8: Stokes's Theorem

We will now show that the curl is related to another vector quantity. Given the vector field:

$$\mathbf{A} = A_x(x,y,z)\,\hat{\mathbf{i}} + A_y(x,y,z)\,\hat{\mathbf{j}} + A_z(x,y,z)\,\hat{\mathbf{k}} \tag{1-117}$$

we will evaluate the surface integral of its curl on the surface, $U(x,y,z) = 0$ for a specified region on that surface. In Figure 1-19 we illustrate the region and we will show that the surface integral of the curl of **A** in that region is related to the line integral of **A** around the perimeter of that region.

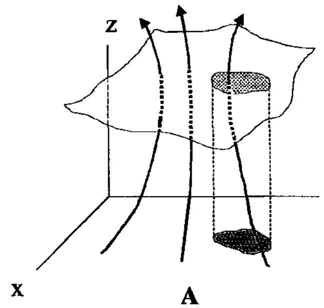


Figure 1-19: The Vector Field A Intersects the Surface U(x,y,z) = 0

First we take the projection of the area of the region onto the x-y plane, as illustrated in Figure 1-19, and then divide this area into infinitesimal increments of area of dimension $\Delta x \Delta y$ as shown in Figure 1-20(a). We will take the line integral about one of the increments as noted in Figure 1-20(b).

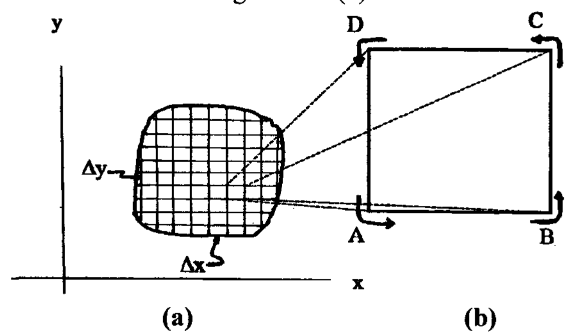


Figure 1-20: (a) The Projection of S on the x-y Plane; (b) an Increment of the Surface has been Enlarged Showing the Direction of Integration

The integral for the incremental area will be given by:

$$\oint_C \mathbf{A} \bullet d\mathbf{l} = \int_A^B A_x\, dx + \int_B^C A_y\, dy + \int_C^D A_x\, dx + \int_D^A A_y\, dy \qquad (1\text{-}118)$$

In this case because the line vector and the field vector lie in the same direction cos $\alpha = 1$ and the line integral and the ordinary integral become the same thing.

Using the mean value theorem the integrals in x are:

$$\int_A^B A_x\, dx \simeq A_x(x_1,y,z)\Delta x \qquad (1\text{-}119)$$

$$\int_C^D A_x dx = -\int_D^C A_x dx \simeq -A_x\,(x_2,y+\Delta y,z)\Delta x \qquad (1\text{-}120)$$

where x_1 and x_2 are mean values of x within the intervals and where the minus sign comes from the fact that the line segment vector points in the negative x direction. Thus the sum of the two integrals is:

$$\int_A^B A_x dx + \int_C^D A_x dx \simeq [A_x(x_1,y,z) - A_x(x_2,y+\Delta y,z)]\,\Delta x \qquad (1\text{-}121)$$

Taking a Taylor Series expansion of the last term in the brackets we have:

$$A_x(x_2,y+\Delta y,z) \simeq A_x(x_2,y,z) + \frac{\partial A_x(x_2,y,z)}{\partial y}\Delta y + \cdots \tag{1-122}$$

and we have:

$$\int_A^B A_x\,dx + \int_C^D A_x dx \simeq \left[A_x(x_1,y,z) - A_x(x_2,y,z) - \frac{\partial A_x(x,y,z)}{\partial y}\Delta y\right] \tag{1-123}$$

as Δx and $\Delta y \to 0$, $x_1 \to x$ and $x_2 \to x$. Thus in this limit:

$$\int_A^B A_x dx + \int_C^D A_x dx \to -\frac{\partial A_x}{\partial y}\Delta x\,\Delta y \tag{1-124}$$

We can go through this process for the other pair of integrals and we will get:

$$\int_B^C A_y\,dy + \int_D^A A_y\,dy \simeq A_y(x+\Delta x,y_1,z)\Delta y - A_y(x,y_2,z)\Delta y \tag{1-125}$$

$$\simeq A_y(x,y_1,z)\Delta y + \frac{\partial A_y(x,y_1,z)}{\partial x}\Delta x\Delta y - A_y(x,y_2,z)\,\Delta y \to \frac{\partial A_y}{\partial x}\Delta x\,\Delta y \tag{1-126}$$

where we have taken the limits $\Delta x \to 0$, $\Delta y \to 0$, meaning that $y_1 \to y$ and $y_2 \to y$. Thus:

$$\oint_{\Delta x\Delta y}\mathbf{A}\bullet d\mathbf{l} = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\Delta x\,\Delta y = (\nabla \times \mathbf{A})_z\,\Delta S_z \tag{1-127}$$

Note that we have evaluated the integral on one of the squares. If we add up all of the contributions we can see from Figure 1-21 that the integrals along adjacent lines will cancel because the value of the integrals on the common line are equal but have opposite signs. If we add up the integrals along the strip Δx wide the only integrals which are not canceled are the ones on the boundary along $\mathbf{dl}_1$ and $\mathbf{dl}_n$ because on those lines there are no adjacent cells. By adding up the integrals from all the cells we are left with only the line integrals on the boundary of the curve. Thus:

$$\lim_{\Delta x,\Delta y \to 0}\sum_{i=1}^{\infty}\oint_{\Delta x\Delta y}\mathbf{A}\bullet d\mathbf{l}_i = \oint_{C_x}\mathbf{A}\bullet d\mathbf{l} = \int_{S_z}\nabla \times \mathbf{A}\bullet d\mathbf{S} \tag{1-128}$$

Similar expressions will result when the same process is used for the x-z plane and the y-z plane. When added to the above expression we will have the equation:

$$\int_S \nabla \times \mathbf{A}\bullet d\mathbf{S} = \oint_C \mathbf{A}\bullet d\mathbf{l} \tag{1-129}$$

where S and C are the actual surface and line.

This is called ***Stokes's Theorem***. Physically it means that the circulation (i.e. curl) of a vector on the surface which is enclosed by a curve can be determined by evaluating the line integral of that vector on that that curve. Not only does this theorem have wide application in physics it also is a very effective way to evaluate difficult surface integrals. Regardless of the nature of the surfaces the term, $\oint_C \mathbf{A}\bullet d\mathbf{l}$, is the same. Therefore, we are led to the non-intuitive result that $\int \nabla \times \mathbf{A}\bullet d\mathbf{S}$ must have the same value for all surfaces bounded by the curve as long as they are continuous.

We have noted that both the divergence and the curl provide a physical description of any vector field. They give us information about the sources of the field and the behavior of the field lines in any region within the field. One would not be surprised then to know that it can be shown that if the curl and divergence of the field are given then the field is uniquely determined, provided the field is well behaved on the boundary of the region and at infinity. This fact is known as ***Helmholtz's Theorem***. Thus finding the divergence and the curl of a vector field becomes very important, as we will illustrate throughout the text, to the understanding of the properties of vector fields.

1-9: Derivatives of Scalar and Vector Products

Since there are several varieties of vector products there will be combinations of derivatives which one has to take into consideration. The gradient is defined only when it operates on a scalar. Thus it can operate only upon the product of two scalars or upon the vector dot product. The gradient of the product of two scalars U and V is:

$$\nabla UV = U\nabla V + V\nabla U \tag{1-130}$$

and the gradient of the dot product of two vectors **A** and **B** is:

$$\nabla(\mathbf{A}\bullet\mathbf{B}) = \mathbf{A}\times(\nabla\times\mathbf{B}) + \mathbf{B}\times(\nabla\times\mathbf{A}) + (\mathbf{A}\bullet\nabla)\mathbf{B} + (\mathbf{B}\bullet\nabla)\mathbf{A} \tag{1-131}$$

The divergence is defined only when it operates upon a vector and thus can be used only with the vector product or the product of a vector and a scalar. For these two situations we have:

$$\nabla\bullet(U\mathbf{A}) = U(\nabla\bullet\mathbf{A}) + \mathbf{A}\bullet\nabla U \tag{1-132}$$

$$\nabla\bullet(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\bullet\nabla\times\mathbf{A} - \mathbf{A}\bullet\nabla\times\mathbf{B} \tag{1-133}$$

Similarly the curl is defined only when it operates on a vector. Thus, it can be used only with a scalar times a vector or with the cross-product. In these two situations we have:

$$\nabla\times(U\mathbf{A}) = U(\nabla\times\mathbf{A}) - \mathbf{A}\times\nabla U \tag{1-134}$$

$$\nabla\times(\mathbf{A}\times\mathbf{B}) = (\mathbf{B}\bullet\nabla)\mathbf{A} - (\mathbf{A}\bullet\nabla)\mathbf{B} + \mathbf{A}(\nabla\bullet\mathbf{B}) - \mathbf{B}(\nabla\bullet\mathbf{A}) \tag{1-135}$$

These relationships can be derived in a straightforward manner by using the vector operations as they have been defined. Note, however, that the rules for vector derivatives are quite different from the analogous scalar derivatives.

1-10: Higher Derivatives

We need to consider also the many ways we can arrange the gradient, the divergence and the curl to make second derivatives. For those operations which are defined all have important applications in physics. It turns out that it is not necessary to go to third derivatives.

The possible second order derivatives are as follows:

(a) **Gradient of the gradient** -- $\nabla(\nabla U)$

This is not defined because the gradient can operate only on a scalar.

(b) **Gradient of the divergence** -- $\nabla(\nabla\bullet\mathbf{A})$

This is defined through the operations already defined above (Equation 1-31):

$$\nabla(\nabla\bullet\mathbf{A}) = \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right)\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right) \tag{1-136}$$

(c) **Gradient of the curl** -- $\nabla(\nabla \times \mathbf{A})$

This is also not defined because the gradient can only operate on a scalar.

(d) **Divergence of the gradient** --$\nabla \bullet \nabla U$ (usually written as $\nabla^2 U$)

Using our definitions of the gradient and the divergence operations we have:

$$\nabla \bullet \nabla U = \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right) \bullet \left(\frac{\partial U}{\partial x}\hat{\mathbf{i}} + \frac{\partial U}{\partial y}\hat{\mathbf{j}} + \frac{\partial U}{\partial z}\hat{\mathbf{k}}\right) \tag{1-137}$$

$$\nabla \bullet \nabla U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = \nabla^2 U \tag{1-138}$$

The operator, ∇^2, is an important differential operator of physics and comes up often enough so that it has been given a special name. It is known as the ***Laplacian*** operator.

(e) **Divergence of the divergence** -- $\nabla \bullet (\nabla \bullet \mathbf{A})$

This is not defined because the divergence can only operate on a vector.

(f) **Divergence of the curl** -- $\nabla \bullet \nabla \times \mathbf{A}$

Using the operations we have defined above:

$$\nabla \bullet \nabla \times \mathbf{A} = \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right) \bullet \left[\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial A_x}{\partial z} + \frac{\partial A_z}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial A_y}{\partial x} + \frac{\partial A_x}{\partial y}\right)\hat{\mathbf{k}}\right] \tag{1-139}$$

When written out this becomes:

$$\nabla \bullet \nabla \times \mathbf{A} = \frac{\partial^2 A_z}{\partial x \partial y} - \frac{\partial^2 A_y}{\partial x \partial z} + \frac{\partial^2 A_x}{\partial y \partial z} - \frac{\partial^2 A_z}{\partial y \partial x} + \frac{\partial^2 A_y}{\partial z \partial x} - \frac{\partial^2 A_x}{\partial z \partial y} \equiv 0 \tag{1-140}$$

Here we have used the fact that the order of differentiation makes no difference. Thus all terms cancel out regardless of the nature of **A** as long as it is a continuous function. This important identity is useful in simplifying vector expressions and we will have occasion to use it often.

(g) **Curl of the gradient** -- $\nabla \times \nabla U$

Using the determinant representation of the curl we have:

$$\nabla \times \nabla U = \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z} \end{bmatrix} \tag{1-141}$$

$$\nabla \times \nabla U = \hat{\mathbf{i}}\left(\frac{\partial^2 U}{\partial y \partial z} - \frac{\partial^2 U}{\partial z \partial y}\right) + \hat{\mathbf{j}}\left(\frac{\partial^2 U}{\partial z \partial x} - \frac{\partial^2 U}{\partial x \partial z}\right) + \hat{\mathbf{k}}\left(\frac{\partial^2 U}{\partial x \partial y} - \frac{\partial^2 U}{\partial y \partial x}\right) \equiv 0 \tag{1-142}$$

where we have again used the fact that the order of differentiation makes no difference. This is another important identity used to simplify vector expressions.

(h) **Curl of the divergence** -- $\nabla \times (\nabla \bullet \mathbf{A})$

This is not defined because the curl can only operate on a vector.

(i) **Curl of the curl** -- $\nabla \times \nabla \times \mathbf{A}$

This is determined by taking the curl twice.

(j) **The Laplacian of a vector.**

This is defined by the relationship:

$$\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \bullet \mathbf{A}) - \nabla^2 \mathbf{A} \tag{1-143}$$

By rearranging terms we have:

$$\nabla^2 \mathbf{A} = \nabla\ (\nabla \bullet \mathbf{A}) - \nabla \times \nabla \times \mathbf{A} \tag{1-144}$$

The terms on the right are defined in (b) and (i) above and are determined by just carrying out the operations. The resultant term represents the Laplacian of the vector. It should be noted and emphasized that only in Cartesian coordinates is the Laplacian of the vector given simply by the Laplacian of the components of the vector. Thus:

$$\nabla^2 \mathbf{A}(x,y,z) = [\nabla^2 A_x(x,y,z)]\,\hat{\mathbf{i}} + [\nabla^2 A_y(x,y,z)]\,\hat{\mathbf{j}} + [\nabla^2 A_z(x,y,z)]\,\hat{\mathbf{k}} \tag{1-145}$$

but in polar coordinates (defined in the next section), for example:

$$\nabla^2 \mathbf{A}(r,\theta,\phi) \neq [\nabla^2 A_r(r,\theta,\phi)]\,\hat{\mathbf{e}}_r + [\nabla^2 A_\theta(r,\theta,\phi)]\,\hat{\mathbf{e}}_\theta + [\nabla^2 A_\phi(r,\theta,\phi)]\,\hat{\mathbf{e}}_\phi \tag{1-146}$$

because, as we will show below, the derivatives of $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\phi$ are not zero.

Problem 1-3

Find the expressions for $\nabla(\nabla \bullet \mathbf{A})$ and $\nabla \times (\nabla \times \mathbf{A})$ in Cartesian coordinates.

1-11: Curvilinear Coordinates

The symmetry of a given problem may be such that Cartesian coordinates would lead to an overly complicated expression or an unusual solution. Indeed the problem may not have solutions which can be expressed in closed form. Changing the relevant equations to another coordinate system may lead to a significant simplification of the problem. Therefore we would like to find out what form basic equations have in other coordinate systems.

In addition to making our problem simpler it is also useful when the equation, $\nabla^2\psi + k^2\psi = 0$, when transformed to another coordinate system, makes the equation separable. There are eleven different coordinate systems for which this is true, elliptic and parabolic coordinate systems, for example, but, for the situations we will encounter in this text, only two coordinate systems in addition to Cartesian coordinates are needed, namely cylindrical coordinates and spherical coordinates.

For Cartesian coordinates we have chosen our coordinates to be x, y and z and our base vectors, as we have noted, are the unit vectors, $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$, which lie on the x, y, and z axes respectively. We define the base vectors in a similar manner for other coordinates. For cylindrical coordinates our coordinates are ρ, ϕ and z and our base vectors are the unit vectors $\hat{\mathbf{e}}_\rho$, $\hat{\mathbf{e}}_\phi$ and $\hat{\mathbf{e}}_z$ which are defined in Figure 1-21. For spherical coordinates our coordinates are r, θ and ϕ and our base vectors are the unit vectors $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\phi$ which are defined in Figure 1-22. We have made these choices for our coordinates so that there will be a minimum of ambiguity between them. The variables in equations will always indicate what coordinate system they have been referenced to. Although some other texts define the base vectors differently ours is by far the most common convention but the student should be aware that there are other conventions in current use.

We will first study the cylindrical coordinate system. In Figure 1-21 we have

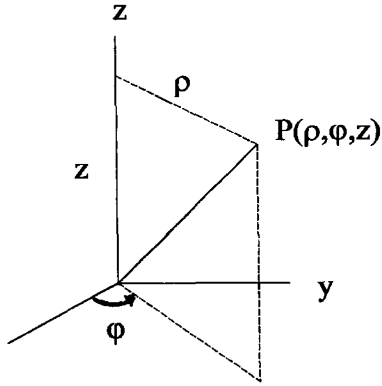


Figure 1-21: Geometric Description of Cylindrical Coordinates

illustrated the way the variables and base vectors are defined geometrically. In terms of the Cartesian coordinates the cylindrical coordinates are:

$$\rho = \sqrt{x^2 + y^2} \tag{1-147}$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{1-148}$$

$$z = z \tag{1-149}$$

The reverse transformation is given by:

$$x = \rho \cos \phi \tag{1-150}$$

$$y = \rho \sin \phi \tag{1-151}$$

$$z = z \tag{1-152}$$

The unit vectors, expressed in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are:

$$\hat{\mathbf{e}}_\rho = \cos \phi\, \hat{\mathbf{i}} + \sin \phi\, \hat{\mathbf{j}} \qquad \hat{\mathbf{i}} = \cos \phi\, \hat{\mathbf{e}}_\rho - \sin \phi\, \hat{\mathbf{e}}_\phi \tag{1-153}$$

$$\hat{\mathbf{e}}_\phi = -\sin \phi\, \hat{\mathbf{i}} + \cos \phi\, \hat{\mathbf{j}} \qquad \hat{\mathbf{j}} = \sin \phi\, \hat{\mathbf{e}}_\rho + \cos \phi\, \hat{\mathbf{e}}_\phi \tag{1-154}$$

$$\hat{\mathbf{e}}_z = \hat{\mathbf{k}} \qquad \hat{\mathbf{k}} = \hat{\mathbf{e}}_z \tag{1-155}$$

One of the great advantages of the Cartesian coordinates is that the base vectors are constant vectors. Note that some of the base vectors in cylindrical coordinates are not constant. For example,

$$\frac{d\hat{\mathbf{e}}_\rho}{d\phi} = -\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}} = \hat{\mathbf{e}}_\phi; \quad \frac{d\hat{\mathbf{e}}_\phi}{d\phi} = -\cos \phi \hat{\mathbf{i}} - \sin \phi; \quad \hat{\mathbf{j}} = -\hat{\mathbf{e}}_\rho \tag{1-156}$$

However all other derivatives of the unit vectors are zero.

The element of length in cylindrical coordinates is:

$$d\mathbf{l} = d\rho\, \hat{\mathbf{e}}_\rho + \rho\, d\phi\, \hat{\mathbf{e}}_\phi + dz\, \hat{\mathbf{e}}_z \tag{1-157}$$

The elements of area are:

$$dA_\rho = \rho\, d\phi\, dz \tag{1-158}$$

$$dA_\phi = d\rho\, dz \tag{1-159}$$

$$dA_z = \rho\, d\rho\, d\phi \tag{1-160}$$

and the element of volume is:

$$d\tau = \rho\, d\rho\, d\phi\, dz \tag{1-161}$$

The expressions for the vector derivative operators are:

The gradient:

$$\nabla U = \frac{\partial U}{\partial \rho}\,\hat{\mathbf{e}}_\rho + \frac{1}{\rho}\frac{\partial U}{\partial \phi}\,\hat{\mathbf{e}}_\phi + \frac{\partial U}{\partial z}\,\hat{\mathbf{e}}_z \tag{1-162}$$

The divergence:

$$\nabla\cdot\mathbf{A} = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho A_\rho) + \frac{1}{\rho}\frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z} \tag{1-163}$$

$$= \frac{\partial A_\rho}{\partial \rho} + \frac{1}{\rho}A_\rho + \frac{1}{\rho}\frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z} \tag{1-164}$$

The curl:

$$\nabla\times\mathbf{A} = \left(\frac{1}{\rho}\frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z}\right)\hat{\mathbf{e}}_\rho + \left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho}\right)\hat{\mathbf{e}}_\phi + \frac{1}{\rho}\left(\frac{\partial}{\partial \rho}(\rho A_\phi) - \frac{\partial A_\rho}{\partial \phi}\right)\hat{\mathbf{e}}_z \tag{1-165}$$

which can also be represented by the determinant:

$$\nabla\times\mathbf{A} = \frac{1}{\rho}\begin{vmatrix} \hat{\mathbf{e}}_\rho & \rho\,\hat{\mathbf{e}}_\phi & \hat{\mathbf{e}}_z \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \phi} & \frac{\partial}{\partial z} \\ A_\rho & \rho\, A_\phi & A_z \end{vmatrix} \tag{1-166}$$

The Laplacian is given by:

$$\nabla^2 U = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\frac{\partial U}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 U}{\partial \phi^2} + \frac{\partial^2 U}{\partial z^2} \tag{1-167}$$

The expressions above can be derived in many ways and a more thorough discussion can be found in the references noted at the end of the chapter. One method is to use the chain rule for derivatives. Thus, in terms of ρ, ϕ and z, the partial of x is:

$$\frac{\partial}{\partial x} = \frac{\partial \rho}{\partial x}\frac{\partial}{\partial \rho} + \frac{\partial \phi}{\partial x}\frac{\partial}{\partial \phi} + \frac{\partial z}{\partial x}\frac{\partial}{\partial z} \tag{1-168}$$

since $\rho = \sqrt{x^2 + y^2}$ we have

$$\frac{\partial \rho}{\partial x} = \frac{x}{(x^2 + y^2)^{1/2}} = \frac{x}{\rho} = \frac{\rho\cos\phi}{\rho} = \cos\phi \tag{1-169}$$

since $\phi = \tan^{-1}\left(\frac{y}{x}\right)$ we have

$$\frac{\partial \phi}{\partial x} = -\left(\frac{y}{x^2}\right)\frac{1}{1 + (y/x)^2} = -\frac{y}{x^2 + y^2} = \frac{\rho\sin\phi}{\rho^2} = -\frac{\sin\phi}{\rho} \tag{1-170}$$

also;

$$\frac{\partial z}{\partial x} = 0 \tag{1-171}$$

Therefore:

$$\frac{\partial}{\partial x} = \cos\phi\,\frac{\partial}{\partial \rho} - \frac{\sin\phi}{\rho}\frac{\partial}{\partial \phi} \tag{1-172}$$

Similarly:

$$\frac{\partial}{\partial y} = \sin\phi\,\frac{\partial}{\partial \rho} + \frac{\cos\phi}{\rho}\frac{\partial}{\partial \phi} \tag{1-173}$$

and

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial z} \tag{1-174}$$

Therefore from these expressions the gradient,

$$\nabla U = \frac{\partial U}{\partial x}\hat{\mathbf{i}} + \frac{\partial U}{\partial y}\hat{\mathbf{j}} + \frac{\partial U}{\partial z}\hat{\mathbf{k}} \tag{1-175}$$

becomes:

$$\nabla U = \left[\left(\cos\phi\frac{\partial}{\partial\rho} - \frac{\sin\phi}{\rho}\frac{\partial}{\partial\phi}\right)U\right](\hat{\mathbf{e}}_\rho\cos\phi - \hat{\mathbf{e}}_\phi\sin\phi)$$

$$+\left[\left(\sin\phi\frac{\partial}{\partial\rho} + \frac{\cos\phi}{\rho}\frac{\partial}{\partial\phi}\right)U\right](\hat{\mathbf{e}}_\rho\sin\phi + \hat{\mathbf{e}}_\phi\cos\phi) + \frac{\partial U}{\partial z}\hat{\mathbf{e}}_z \tag{1-176}$$

which reduces to Equation 1-164 above. A similar process will yield the other equations.

Spherical coordinates, whose geometry is illustrated in Figure 1-22, are defined as:

$$r = \sqrt{x^2 + y^2 + z^2} \tag{1-177}$$

$$\theta = \cos^{-1}\left[\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right] \tag{1-178}$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{1-179}$$

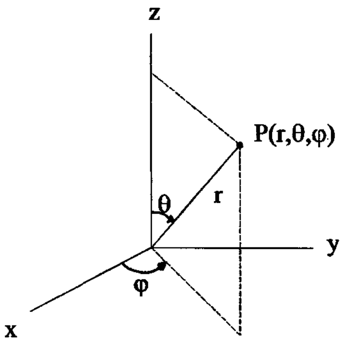


Figure 1-22: Geometric Description of Spherical Coordinates

Conversely the Cartesian coordinates in terms of the spherical coordinates are:

$$x = r\sin\theta\cos\phi \tag{1-180}$$

$$y = r\sin\theta\sin\phi \tag{1-181}$$

$$z = r\cos\theta \tag{1-182}$$

The unit vectors are:

$$\hat{\mathbf{e}}_r = \sin\theta\cos\phi\,\hat{\mathbf{i}} + \sin\theta\sin\phi\,\hat{\mathbf{j}} + \cos\theta\,\hat{\mathbf{k}}$$

$$\hat{\mathbf{i}} = \sin\theta\cos\phi\,\hat{\mathbf{e}}_r + \cos\theta\cos\phi\,\hat{\mathbf{e}}_\theta - \sin\phi\,\hat{\mathbf{e}}_\phi \tag{1-183}$$

$$\hat{\mathbf{e}}_\theta = \cos\theta\cos\phi\,\hat{\mathbf{i}} + \cos\theta\sin\phi\,\hat{\mathbf{j}} - \sin\theta\,\hat{\mathbf{k}}$$

$$\hat{\mathbf{j}} = \sin\theta \sin\phi\, \hat{\mathbf{e}}_r + \cos\theta \sin\phi\, \hat{\mathbf{e}}_\theta + \cos\phi\, \hat{\mathbf{e}}_\phi \tag{1-184}$$

$$\hat{\mathbf{e}}_\phi = -\sin\phi\, \hat{\mathbf{i}} + \cos\phi\, \hat{\mathbf{j}} \qquad \hat{\mathbf{k}} = \cos\theta\, \hat{\mathbf{e}}_r - \sin\theta\, \hat{\mathbf{e}}_\theta \tag{1-185}$$

Again all of the base vectors are not constant:

$$\frac{d\hat{\mathbf{e}}_r}{dr} = 0 \qquad \frac{d\hat{\mathbf{e}}_\theta}{dr} = 0 \qquad \frac{d\hat{\mathbf{e}}_\phi}{dr} = 0 \tag{1-186}$$

$$\frac{d\hat{\mathbf{e}}_r}{d\theta} = \cos\theta \cos\phi\, \hat{\mathbf{i}} + \cos\theta \sin\phi\, \hat{\mathbf{j}} - \sin\theta\, \hat{\mathbf{k}} = \hat{\mathbf{e}}_\theta \tag{1-187}$$

$$\frac{d\hat{\mathbf{e}}_\theta}{d\theta} = -\sin\theta \cos\phi\, \hat{\mathbf{i}} - \sin\theta \sin\phi\, \hat{\mathbf{j}} - \cos\theta\, \hat{\mathbf{k}} = -\hat{\mathbf{e}}_r \tag{1-188}$$

$$\frac{d\hat{\mathbf{e}}_\phi}{d\theta} = 0 \tag{1-189}$$

$$\frac{d\hat{\mathbf{e}}_r}{d\phi} = -\sin\theta \sin\phi\, \hat{\mathbf{i}} + \sin\theta \cos\phi\, \hat{\mathbf{j}} = \sin\theta\, \hat{\mathbf{e}}_\phi \tag{1-190}$$

$$\frac{d\hat{\mathbf{e}}_\theta}{d\phi} = -\cos\theta \sin\phi\, \hat{\mathbf{i}} + \cos\theta \cos\phi\, \hat{\mathbf{j}} = \cos\theta\, \hat{\mathbf{e}}_\phi \tag{1-191}$$

$$\frac{d\hat{\mathbf{e}}_\phi}{d\phi} = -\cos\phi\, \hat{\mathbf{i}} - \sin\phi\, \hat{\mathbf{j}} = -\sin\theta\, \hat{\mathbf{e}}_r - \cos\theta\, \hat{\mathbf{e}}_\theta \tag{1-192}$$

The element of length in spherical coordinates is:

$$d\mathbf{l} = dr\, \hat{\mathbf{e}}_r + r\, d\theta\, \hat{\mathbf{e}}_\theta + r \sin\theta\, d\phi\, \hat{\mathbf{e}}_\phi \tag{1-193}$$

Similarly the elements of area are:

$$dA_r = r^2 \sin\theta\, d\theta\, d\phi \tag{1-194}$$

$$dA_\theta = r \sin\theta\, d\phi\, dr \tag{1-195}$$

$$dA_\phi = r\, dr\, d\theta \tag{1-196}$$

and the element for volume is:

$$d\tau = r^2 dr \sin\theta\, d\theta\, d\phi \tag{1-197}$$

The vector derivatives are as follows:

The gradient is:

$$\nabla U = \frac{\partial U}{\partial r}\, \hat{\mathbf{e}}_r + \frac{1}{r} \frac{\partial U}{\partial \theta}\, \hat{\mathbf{e}}_\theta + \frac{1}{r \sin\theta} \frac{\partial U}{\partial \phi}\, \hat{\mathbf{e}}_\phi \tag{1-198}$$

The divergence is;

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta} (\sin\theta\, A_\theta) + \frac{1}{r \sin\theta} \frac{\partial A_\phi}{\partial \phi} \tag{1-199}$$

The curl is:

$$\nabla \times \mathbf{A} = \frac{1}{r \sin\theta} \left[\frac{\partial}{\partial \theta} (\sin\theta\, A_\phi) - \frac{\partial A_\theta}{\partial \phi} \right] \hat{\mathbf{e}}_r + \frac{1}{r \sin\theta} \left[\frac{\partial A_r}{\partial \phi} - \sin\theta \frac{\partial}{\partial r} (r A_\phi) \right] \hat{\mathbf{e}}_\theta + \frac{1}{r} \left[\frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial \theta} \right] \hat{\mathbf{e}}_\phi \tag{1-200}$$

Again we can use the determinant notation;

$$\nabla \times \mathbf{A} = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\mathbf{e}}_r & r\, \hat{\mathbf{e}}_\theta & r \sin\theta\, \hat{\mathbf{e}}_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ A_r & r A_\theta & r \sin\theta\, A_\phi \end{vmatrix} \tag{1-201}$$

The Laplacian is;

$$\nabla^2 U = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial U}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 U}{\partial\phi^2} \qquad (1\text{-}202)$$

It is instructive to see how the expressions for certain functions differ according to what coordinate system is being used. In Figure 1-23 we illustrate one way to

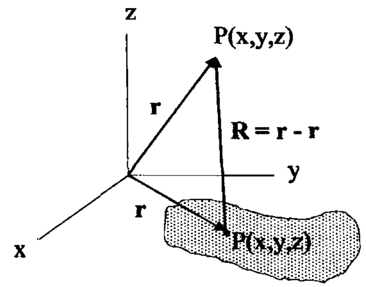


Figure 1-23: The Definition of the Position Vector

express the displacement vector, which is also the convention we will use in this text. Because it often is not convenient or possible to use the origin as the reference, we define the position vector to be:

$$\mathbf{R} = (x - x')\,\hat{\mathbf{i}} + (y - y')\,\hat{\mathbf{j}} + (z - z')\,\hat{\mathbf{k}} \qquad (1\text{-}203)$$

Here the coordinates x,y,z refer to a region where the observer is located who is recording data at a point, P(x,y,z), which we call the field point. The coordinates x', y' and z' refer to a region where the effects being observed are generated from. A point in this region, $P(x',y',z')$ is therefore called a source point. For example, if we were studying the motion of an object in the earth's gravitational field,P(x,y,z) would be a point in the object and $P(x',y',z')$ would be a point on the earth.

We now explore the applications of the vector operator upon this representation of distance. As a first example we find the gradient of the term 1/R. In Cartesian coordinates we have:

$$\frac{1}{R} = \frac{1}{\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}} \qquad (1\text{-}204)$$

The gradient is:

$$\nabla\left(\frac{1}{R}\right) = \frac{\partial}{\partial x}\left(\frac{1}{R}\right)\hat{\mathbf{i}} + \frac{\partial}{\partial y}\left(\frac{1}{R}\right)\hat{\mathbf{j}} + \frac{\partial}{\partial z}\left(\frac{1}{R}\right)\hat{\mathbf{k}} \qquad (1\text{-}205)$$

$$\nabla\left(\frac{1}{R}\right) = \frac{(-1/2)2(x-x')\,\hat{\mathbf{i}} + (-1/2)2(y-y')\,\hat{\mathbf{j}} + (-1/2)2(z-z')\,\hat{\mathbf{k}}}{\left[(x-x')^2+(y-y')^2+(z-z')^2\right]^{3/2}} \qquad (1\text{-}206)$$

$$= -\frac{\mathbf{R}}{R^3} = -\frac{\hat{\mathbf{e}}_R}{R^2} \qquad (1\text{-}207)$$

Note that if we assumed that r was constant and r' varied then:

$$\nabla'\left(\frac{1}{R}\right) = \frac{\hat{\mathbf{e}}_R}{R^2} = -\nabla\left(\frac{1}{R}\right) \qquad (1\text{-}208)$$

where ∇' means differentiation with respect to the primed coordinates.

To use the other coordinate systems we need to take into consideration that the base vectors change as we move from R to R'. They point in one direction along **r** and in another direction along $\mathbf{r}'$ whereas the vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ always point in the same direction.

To find out what $\mathbf{R}$ would look like in the other coordinate systems it is convenient to start with the expression in Cartesian coordinates and use the unit vector transformations Equations 1-153 through 1-155 and Equations 1-183 through 1-185. Thus, since in Cartesian coordinates:

$$\mathbf{R} = (x - x')\,\hat{\mathbf{i}} + (y - y')\,\hat{\mathbf{j}} + (z - z')\,\hat{\mathbf{k}} \tag{1-209}$$

in cylindrical coordinates $\mathbf{R}$ becomes:

$$\mathbf{R} = (\rho\cos\phi - \rho'\cos\phi')(\cos\phi\,\hat{\mathbf{e}}_\rho - \sin\phi\,\hat{\mathbf{e}}_\phi) + (\rho\sin\phi - \rho'\sin\phi')(\sin\phi\,\hat{\mathbf{e}}_\rho + \cos\phi\,\hat{\mathbf{e}}_\phi) + (z - z')\,\hat{\mathbf{e}}_z \tag{1-210}$$

which reduces to:

$$\mathbf{R} = [\rho - \rho'\cos(\phi - \phi')]\,\hat{\mathbf{e}}_\rho + \rho'\sin(\phi - \phi')\,\hat{\mathbf{e}}_\phi + (z - z')\,\hat{\mathbf{e}}_z \tag{1-211}$$

In spherical coordinate the expression for $\mathbf{R}$ is even more complicated:

$$\mathbf{R} = [r - r'\sin\theta\sin\theta'\cos(\phi - \phi') - r'\cos\theta\cos\theta']\,\hat{\mathbf{e}}_r + r'\,[\sin\theta\cos\theta' - \cos\theta\sin\theta'\cos(\phi - \phi')]\,\hat{\mathbf{e}}_\theta + r'\sin\theta'\sin(\phi - \phi')\,\hat{\mathbf{e}}_\phi \tag{1-212}$$

Let us find the divergence of $\mathbf{R}$ in Cartesian coordinates. Then:

$$\nabla\bullet\mathbf{R} = \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right)\bullet\left[(x - x')\,\hat{\mathbf{i}} + (y - y')\,\hat{\mathbf{j}} + (z - z')\,\hat{\mathbf{k}}\right] \tag{1-213}$$

$$= 1 + 1 + 1 = 3 \tag{1-214}$$

Now lets try this with cylindrical coordinates:

$$\nabla\bullet\mathbf{R} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho R_\rho) + \frac{1}{\rho}\frac{\partial R_\phi}{\partial\phi} + \frac{\partial R_z}{\partial z} \tag{1-215}$$

$$= \frac{1}{\rho}\frac{\partial}{\partial\rho}[\rho^2 - \rho\rho'\cos(\phi - \phi')] + \frac{1}{\rho}\frac{\partial}{\partial\phi}[\rho'\sin(\phi - \phi')] + \frac{\partial}{\partial z}(z - z') \tag{1-216}$$

$$= \frac{1}{\rho}[2\rho - \rho'\cos(\phi - \phi')] + \frac{1}{\rho}\rho'\cos(\phi - \phi') + 1 = 2 + 1 = 3 \tag{1-217}$$

Similarly, if we used spherical coordinates we would obtain the same answer.

Problem 1-4

Find the gradient and Laplacian of the following scalar fields:

(a) $U(x,y,z) = x^2 + 2xy + y^2$

(b) $U(x,y,z) = 2xyz$

(c) $U(\rho,\phi,z) = \dfrac{\cos\phi}{\rho^2}$

(d) $U(r,\theta,\phi) = r^2\sin\theta\cos\phi$

Problem 1-5

Show that:

(a) $(\mathbf{A}\bullet\nabla)\mathbf{r} = \mathbf{A}$

(b) $\nabla(\mathbf{A}\bullet\mathbf{r}) = \mathbf{A}$ where $\mathbf{A}$ is a constant vector

(c) $\nabla U(r) = \dfrac{\mathbf{r}}{r}\dfrac{dU}{dr} = \dfrac{dU}{dr}\,\hat{\mathbf{e}}_r$

(d) $\nabla(\mathbf{A}\bullet\mathbf{B}\times\mathbf{r}) = \mathbf{A}\times\mathbf{B}$ where **A** and **B** are constant vectors

Problem 1-6

Find U(r) if $\nabla\bullet U(r)\mathbf{r} = 0$

Problem 1-7

Find the divergence of the following functions:

(a) $\mathbf{A} = x\hat{\mathbf{i}} + y^3\hat{\mathbf{j}} - 2z\hat{\mathbf{k}}$

(b) $\mathbf{A} = x^2\hat{\mathbf{i}} - y^2\hat{\mathbf{j}} + 2xyz\hat{\mathbf{k}}$

(c) $\mathbf{A} = \rho^2\sin\phi\,\hat{\mathbf{e}}_\rho + \cos\phi\,\hat{\mathbf{e}}_\phi + \hat{\mathbf{e}}_z$

(d) $\mathbf{A} = \frac{\sin\phi}{r}\hat{\mathbf{e}}_r - \frac{\cos\phi}{r}\hat{\mathbf{e}}_\theta + r\cos\theta\,\hat{\mathbf{e}}_\phi$

Problem 1-8

Find the curl of the following functions:

(a) $\mathbf{A} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$

(b) $\mathbf{A} = yz\hat{\mathbf{i}} + xz\hat{\mathbf{j}} + xy\hat{\mathbf{k}}$

(c) $\mathbf{A} = \rho\cos\phi\,\hat{\mathbf{e}}_\rho + \rho\sin\phi\,\hat{\mathbf{e}}_\phi + \rho^2\,\hat{\mathbf{e}}_z$

(d) $\mathbf{A} = \cos\phi\,\hat{\mathbf{e}}_r + \sin\phi\,\hat{\mathbf{e}}_\theta + r\cos\theta\,\hat{\mathbf{e}}_\phi$

Problem 1-9

Find $\oint_c \mathbf{A}\bullet d\mathbf{l}$ along any of the paths defined below if $\mathbf{A} = 2xy^2\hat{\mathbf{i}} + 2x^2y\hat{\mathbf{j}}$:

(a) (0,0) → (1,0) → (1,1); on the x-axis and then parallel to the y-axis

(b) (0,0) → (1,1) along $x = y$

(c) (0,0) → (1,1) along $y = x^2$

Problem 1-10

Given the vector field, $\mathbf{A} = xy\,\hat{\mathbf{i}} + 2yz\,\hat{\mathbf{j}} + 3xz\,\hat{\mathbf{k}}$, find its line integral around the triangle (0,0,0) → (0,2,0) → (0,0,2) → (0,0,0). Verify the answer using Stokes's Theorem.

Problem 1-11

Find the line integral of the vector field, $\mathbf{A} = y\hat{\mathbf{i}} - x\hat{\mathbf{j}}$, moving clockwise on the curve $x^2 + y^2 = 1$. Check by using Stokes's Theorem.

Problem 1-12

The vector field, $\mathbf{A} = xy\hat{\mathbf{i}} + 3y\hat{\mathbf{j}} + xz\hat{\mathbf{k}}$, intersects the surface defined by the function $x + y + z = 1$ and the coordinate planes. Find the flux through the surface using the surface integral and check using the divergence theorem.

Problem 1-13

Find the flux of the field, $\mathbf{A} = x\hat{\mathbf{i}} + 2y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$, through the hemisphere given by the equations $x^2 + y^2 + z^2 = 1$, $z \geq 0$, using both the surface integral and the divergence theorem.

Problem 1-14

Show that $\nabla\bullet\mathbf{r} = \nabla\bullet\mathbf{R} = 3$ for spherical coordinates.

1-12: Tensors and Matrices

Many types of physical phenomena require more designation than just a magnitude, such as scalars, or magnitude and direction such as vectors. For example, force in a material applied along the x-axis may also create stresses along the y- and z-axes. We have no way of describing this through the vector description and instead need to invoke the properties of a mathematical entity which we call a **tensor**.

To introduce tensors we will define them by their transformation properties. Thus a tensor is any quantity which transforms from one coordinate system (x, y, z) to another (x′, y′, z′) as:

$$\mathbb{T}'_{ijk\ldots} = \textstyle\sum_i \sum_j \sum_k \cdots C_{il} C_{jm} C_{kn} \cdots \mathbb{T}_{lmn\ldots} \qquad (1\text{-}218)$$

The number of subscripts defines the rank of the tensor. For example, for a tensor of rank zero we have:

$$\mathbb{T}' = \mathbb{T} \qquad (1\text{-}219)$$

which means that there is no transformation at all. It is a pure number and thus a scalar.

A tensor of rank one transforms as:

$$\mathbb{T}'_i = \textstyle\sum_j C_{ij}\, \mathbb{T}_j \qquad (1\text{-}220)$$

We will see that vectors transform in this way and thus vectors are technically tensors of rank one.

A tensor of rank two transforms as:

$$\mathbb{T}'_{ij} = \textstyle\sum_l \sum_m C_{il} C_{jm} \mathbb{T}_{lm} \qquad (1\text{-}221)$$

Tensor interactions are fairly common. In addition to stress, which we have already mentioned, the quantities of moment of inertia and dielectric polarization and index of refraction require tensors to describe their properties most generally. Linear or vector approximations are usually made, however. They are easier to deal with.

Usually, a tensor of rank two is adequate to describe most natural phenomena. Thus the term tensor usually refers to a second rank tensor.

A common method of working with tensors involves using matrices to represent the transformation equations. We will therefore summarize the properties of matrices and show how they are applied to tensors.

Simply defined a matrix is an array of numbers which are arranged in rows and columns and which obey certain rules of algebra. We represent a matrix by writing the array in brackets. For example, a general matrix, **a**, is written as:

$$\mathbf{a} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad (1\text{-}222)$$

The entity, **a**, is referred to as an mxn (m by n) matrix or a matrix of size m by n. The elements of the matrix are the a_{ij} and by convention i refers to the row and j to the column.

If we have two matrices, **a** and **b**, the sum, **c**, is defined to be the matrix whose elements are $c_{ij} = a_{ij} + b_{ij}$. This means that matrix addition has meaning only if **a** and **b** are the same size. Thus **c** = **a** + **b** is written as:

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \quad (1\text{-}223)$$

From this definition we can see that the matrices obey the commutative and associative laws of addition. Thus:

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad (1\text{-}224)$$

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c} \quad (1\text{-}225)$$

Subtraction is defined in the same way. Thus if **c** = **a** − **b** then $c_{ij} = a_{ij} - b_{ij}$.

Multiplication by a scalar **c** = k**a** means that each element of the matrix is multiplied by the scalar, k, and the elements of **c** are $c_{ij} = ka_{ij}$.

The product of **a** and **b**, **c** = **ab**, has components defined as:

$$c_{ij} = \textstyle\sum_k a_{ik} b_{kj} \quad (1\text{-}226)$$

The matrices **a** and **b** do not have to be the same size. However the number of columns of **a** has to be equal to the number of rows of **b** or the product of the two matrices is not defined. This means also that **ab** is not necessarily equal to **ba**. It is possible that one product exists but not the other. To illustrate the process of multiplication we take the two general 2 × 2 matrices, **a** and **b**:

$$\mathbf{a} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \quad (1\text{-}227)$$

The different products are:

$$\mathbf{ab} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \quad (1\text{-}228)$$

$$\mathbf{ba} = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix} \quad (1\text{-}229)$$

As we can see, in general, no element of either matrix is the same. [Multiplication of matrices can be remembered by noting that the product c_{11} is found by multiplying each element of the first row of **a** by the corresponding element of the first column of **b**. Similarly, c_{21} is the product of each element of the second row of **a** by the corresponding element of the first column of **b** and so on until each element of **c** has been found.]

A unit matrix can be defined such that:

$$\mathbf{1a} = \mathbf{a1} \quad (1\text{-}230)$$

and will be valid for any matrix **a** . The elements of the matrix are

$$1_{ij} = \delta_{ij} \tag{1-231}$$

where δ_{ij} is the Kronecker δ,

$$\delta_{ij} = 0 \qquad \text{if } i \neq j \tag{1-232}$$

$$\delta_{ij} = 1 \qquad \text{if } i = j \tag{1-233}$$

Thus the unit matrix is represented as:

$$\mathbf{1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{1-234}$$

To indicate how matrices are used in working with the transformation of tensors, we use the second order tensor, $\mathbb{T}_{ij}$.

As we noted the transformation is defined as:

$$\mathbb{T}'_{ij} = \sum_l \sum_m C_{il} C_{jm} \mathbb{T}_{lm} \tag{1-235}$$

We note that the transformation equation is not consistent with the definition of matrix multiplication, Equation 1-226. To put it in a form so that matrix techniques can be used we rearrange the terms:

$$\mathbb{T}'_{ij} = \sum_l \sum_m C_{il} C_{jm} \mathbb{T}_{lm} = \sum_l \sum_m C_{il} \mathbb{T}_{lm} C_{jm} = \sum_l \sum_m C_{il} \mathbb{T}_{lm} \widetilde{C}_{mj} \tag{1-236}$$

where $\widetilde{C}_{mj}$ is the transpose of the matrix **c** and is created by exchanging the rows by the columns. For example, for the two-by-two matrix:

$$\mathbf{a} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \tag{1-237}$$

the transpose is:

$$\tilde{\mathbf{a}} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} \tag{1-238}$$

Written in matrix notation, the transformation of a tensor can therefore be given by:

$$\mathbb{T}' = \mathbf{a}\mathbb{T}\tilde{\mathbf{a}} \tag{1-239}$$

where the tensor itself is represented as a matrix.

For a certain class of transformations the matrices representing the transformations have some very useful properties. For simplicity we will restrict ourselves to tensors of the first rank or vectors, although what we will prove will apply to tensors of any rank. Because physical entities represented by vectors, such as force and position, do not change their magnitudes under a coordinate transformation, we require that the length of our general vector remain invariant after all transformations. The vector transformations are represented by the two equations:

$$r'_i = \Sigma_j a_{ij} r_j \tag{1-240}$$

$$r_i = \Sigma_k b_{ik} r'_k \tag{1-241}$$

and, to keep the length of the vector constant, they must obey the equation:

$$\Sigma_i (r_i')^2 = \Sigma_j \left(r_j\right)^2 \tag{1-242}$$

We transform the left side getting:

$$\Sigma_i \left(r_i'\right)^2 = \Sigma_i \Sigma_j\ a_{ij} r_j \Sigma_k\ a_{ik} r_k = \Sigma_i \Sigma_j \Sigma_k\ a_{ij} a_{ik} r_j r_k = \Sigma_j\ \left(r_j\right)^2 \tag{1-243}$$

which is true only if:

$$\Sigma_i a_{ij} a_{ik} = \delta_{jk} \tag{1-244}$$

then:

$$\Sigma_i \Sigma_j \Sigma_k\ a_{ij} a_{ik} r_j r_k = \Sigma_j \Sigma_k \delta_{jk} r_j r_k = \Sigma_j\ \left(r_j\right)^2 \tag{1-245}$$

Similarly if we had started with $(r_j)^2$ we would have ended up with:

$$\Sigma_i b_{ij} b_{ik} = \delta_{jk} \tag{1-246}$$

We can find other properties of the transformation matrix by requiring that the inverse transformation return r_i and r_i' back into themselves. Thus:

$$r_i' = \Sigma_j a_{ij} r_j = \Sigma_j \Sigma_k a_{ij} b_{jk} r_k' \tag{1-247}$$

is true only if:

$$\Sigma_j a_{ij} b_{jk} = \delta_{ik} \tag{1-248}$$

This is in the form of matrix multiplication and says that:

$$\mathbf{ab} = \mathbf{1} \tag{1-249}$$

which in turn means that:

$$\mathbf{b} = \mathbf{a}^{-1} \tag{1-250}$$

If we had started with r_i we would have ended up with the equation:

$$\Sigma_j b_{ij} a_{jk} = \delta_{ik} \tag{1-251}$$

which leads to the same results.

Finally we find another relationship by considering the product:

$$\Sigma_i \Sigma_k a_{kp} a_{ki} b_{ij} = \Sigma_i \delta_{pi} b_{ij} = b_{pj} \tag{1-252}$$

where we have used Equation 1-244. However the product is also:

$$\Sigma_i \Sigma_k a_{kp} a_{ki} b_{ij} = \Sigma_k a_{kp} \delta_{kj} = a_{jp} \tag{1-253}$$

where we have used Equation 1-248. This means that:

$$b_{pj} = a_{jp} = \tilde{a}_{pj} \tag{1-254}$$

or in matrix notation:

$$\mathbf{b} = \tilde{\mathbf{a}} \tag{1-255}$$

but we have already noted that $\mathbf{b} = \mathbf{a}^{-1}$ thus:

$$\mathbf{a}^{-1} = \tilde{\mathbf{a}} \tag{1-256}$$

This means that the inverse transformation can be found just by converting the transformation matrix to its transpose. Furthermore the product of the transformation matrix with the inverse transformation matrix is the unit matrix, **1**. Transformations which have these properties are known as orthonormal transformations and the matrices which represent them are called orthogonal matrices.

We now look at vectors once more in the context of their transformation properties and demonstrate some properties of their transformation matrices.

Suppose we have two orthogonal coordinate systems with a common origin but whose axes are not aligned. One coordinate system will have coordinates (x,y,z) and the other coordinate system will have coordinates (x′,y′z′). The base vectors ($\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$) and ($\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, $\hat{\mathbf{k}}'$) are related by the transformation equation:

$$\hat{\mathbf{r}}' = \mathbf{a}\hat{\mathbf{r}} \tag{1-257}$$

which we can alternatively write as:

$$r_i' = \Sigma_j a_{ij} r_j \tag{1-258}$$

In matrix notation the relationship between $\hat{\mathbf{i}}'$, $\hat{\mathbf{j}}'$, $\hat{\mathbf{k}}'$ and $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$, and thus (x′,y′,z′) and (x,y,z), is:

$$\begin{bmatrix} \hat{\mathbf{i}}' \\ \hat{\mathbf{j}}' \\ \hat{\mathbf{k}}' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} \tag{1-259}$$

or as the set of equations:

$$\hat{\mathbf{i}}' = a_{11}\hat{\mathbf{i}} + a_{12}\hat{\mathbf{j}} + a_{13}\hat{\mathbf{k}} \tag{1-260}$$

$$\hat{\mathbf{j}}' = a_{21}\hat{\mathbf{i}} + a_{22}\hat{\mathbf{j}} + a_{23}\hat{\mathbf{k}} \tag{1-261}$$

$$\hat{\mathbf{k}}' = a_{31}\hat{\mathbf{i}} + a_{32}\hat{\mathbf{j}} + a_{33}\hat{\mathbf{k}} \tag{1-262}$$

We know that the inverse transformation is:

$$\mathbf{r} = \mathbf{b}\mathbf{r}' \tag{1-263}$$

which we can write as:

$$r_i = \Sigma_j b_{ij} r_j' = \Sigma_j \tilde{a}_{ij} r_j' \tag{1-264}$$

which means that the transformation for the base vectors is:

$$\hat{\mathbf{r}} = \tilde{\mathbf{a}}\,\hat{\mathbf{r}}' \tag{1-265}$$

which we write in matrix form as:

$$\begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{i}}' \\ \hat{\mathbf{j}}' \\ \hat{\mathbf{k}}' \end{bmatrix} \tag{1-266}$$

and when written out is:

$$\hat{\mathbf{i}} = a_{11}\hat{\mathbf{i}}' + a_{21}\hat{\mathbf{j}}' + a_{31}\hat{\mathbf{k}}' \tag{1-267}$$

$$\hat{\mathbf{j}} = a_{12}\hat{\mathbf{i}}' + a_{22}\hat{\mathbf{j}}' + a_{32}\hat{\mathbf{k}}' \tag{1-268}$$

$$\hat{\mathbf{k}} = a_{13}\hat{\mathbf{i}}' + a_{23}\hat{\mathbf{j}}' + a_{33}\hat{\mathbf{k}}' \tag{1-269}$$

Furthermore, by taking the appropriate dot product, the components of **a** and $\tilde{\mathbf{a}}$ are seen to be:

$$\begin{array}{lll} \hat{\mathbf{i}}\cdot\hat{\mathbf{i}}' = a_{11} & \hat{\mathbf{i}}\cdot\hat{\mathbf{j}}' = a_{21} & \hat{\mathbf{i}}\cdot\hat{\mathbf{k}}' = a_{31} \\ \hat{\mathbf{j}}\cdot\hat{\mathbf{i}}' = a_{12} & \hat{\mathbf{j}}\cdot\hat{\mathbf{j}}' = a_{22} & \hat{\mathbf{j}}\cdot\hat{\mathbf{k}}' = a_{32} \\ \hat{\mathbf{k}}\cdot\hat{\mathbf{i}}' = a_{13} & \hat{\mathbf{k}}\cdot\hat{\mathbf{j}}' = a_{23} & \hat{\mathbf{k}}\cdot\hat{\mathbf{k}}' = a_{33} \end{array} \tag{1-270}$$

Because our transformation is orthogonal the matrices **a** and $\tilde{\mathbf{a}}$ obey the equations:

$$\mathbf{a}\tilde{\mathbf{a}} = \mathbf{1} \qquad (1\text{-}271)$$

and:

$$\tilde{\mathbf{a}} = \mathbf{a}^{-1} \qquad (1\text{-}272)$$

We will now apply this to the situation of a vector transformation due to a simple rotation around the z-axis of angle α. In Figure 1-26 we illustrate the components (A_x and A_y) of the vector **A**, in the coordinate system x,y and the components of the vector **A′** ($A_{x'}$ and $A_{y'}$) in the coordinate system x′,y′.

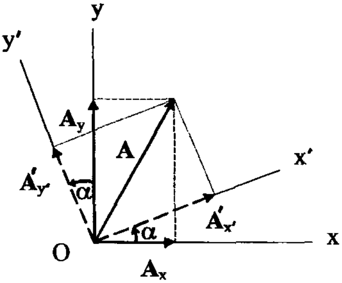


Figure 1-24: Rotation of a Vector about the z-Axis

In x,y our vector is represented by:

$$\mathbf{A} = A_x\hat{\mathbf{i}} + A_y\hat{\mathbf{j}} \qquad (1\text{-}273)$$

In x′,y′ the vector is represented by:

$$\mathbf{A}' = A'_{x'}\hat{\mathbf{i}}' + A'_{y'}\hat{\mathbf{j}}' \qquad (1\text{-}274)$$

From Figure 1-26 we can see that

$$A'_{x'} = A_x \cos\alpha + A_y \sin\alpha \qquad (1\text{-}275)$$

$$A'_{y'} = -A_x \sin\alpha + A_y \cos\alpha \qquad (1\text{-}276)$$

Thus:

$$\begin{bmatrix} A'_{x'} \\ A'_{y'} \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} A_x \\ A_y \end{bmatrix} \qquad (1\text{-}277)$$

or in shorthand notation:

$$\mathbf{A}' = \mathbf{R}\,\mathbf{A} \qquad (1\text{-}278)$$

where the matrix **R** is the 2×2 rotation matrix of Equation 1-277.

Also from Figure 1-26 we can see that the reverse transformation is:

$$A_x = A'_{x'} \cos\alpha - A'_{y'} \sin\alpha \qquad (1\text{-}279)$$

$$A_y = A'_{x'} \sin\alpha + A'_{y'} \cos\alpha \qquad (1\text{-}280)$$

which in matrix form is:

$$\begin{bmatrix} A_x \\ A_y \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} A'_{x'} \\ A'_{y'} \end{bmatrix} \qquad (1\text{-}281)$$

or in our shorthand notation:

$$\mathbf{A} = \tilde{\mathbf{R}}\mathbf{A}' \qquad (1\text{-}282)$$

Thus, from looking just at the geometry of the transformation, we see that a vector transforms like a tensor. Note that:

$$\tilde{\mathbf{R}}\mathbf{R} = \mathbf{R}^{-1}\mathbf{R} = \mathbf{1}. \tag{1-283}$$

So far we have emphasized the transformation properties of tensors because that provides the basis for their mathematical definition. It also provides a way to demonstrate how the mathematics of matrices can be used in tensor algebra. For tensors we define the dot product (often called as the inner product). In summation notation the dot product can be written as:

$$C_{jk} = \Sigma_i A_{ji} B_{ik} \tag{1-284}$$

which is simply matrix multiplication of the tensors represented as matrices. The divergence, the dot product with the vector operator ∇, also is defined as is the curl, the cross products with the vector operator ∇. For example, the divergence is given as:

$$\nabla \bullet \mathbb{T} = \frac{\partial T_{ij}}{\partial x_i} \hat{\mathbf{e}}_{ij} \tag{1-285}$$

A well-known example of a tensor dot product is the determination of angular momentum from the moment of inertia tensor. The angular momentum, **L**, the angular acceleration, $\boldsymbol{\omega}$, and the moment of inertia, $\mathbb{I}$, obey the relationship

$$\mathbf{L} = \mathbb{I} \bullet \boldsymbol{\omega} \tag{1-286}$$

which is a shorthand way of writing

$$L_i = \sum_j I_{ij} \omega_j \tag{1-287}$$

and therefore each component will be given by:

$$L_x = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \tag{1-288}$$

$$L_y = I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \tag{1-289}$$

$$L_z = I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z \tag{1-290}$$

In matrix notation we write this as:

$$\begin{bmatrix} L_x \\ L_y \\ L_z \end{bmatrix} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{1-291}$$

This set of equations is a good example of what a tensor does. Note that the moment of inertia tensor, $\mathbb{I}$, causes all of the components of ω to mix with the component L_x. For an interaction which involves just vectors a component which is at right angles to the vector cannot change the vector. For example, we would be surprised if we pushed an object to the north and it instead went west. However, for a tensor interaction this is possible. When we discussed the cross product, $\mathbf{A} \times \mathbf{B}$, we noted that it created a vector at right angles to both **A** and **B** and we noted that it did not transform in coordinate inversions like a vector. It is actually a tensor.

So far we have restricted ourselves to Cartesian Tensors. The extension to a general coordinate system is straightforward but adds some complications. For example, we know from the properties of partial derivatives that the derivative of the Cartesian coordinates x'_i in terms of the coordinates x_i is:

$$dx'_i = \sum_j \frac{\partial x'_i}{\partial x_j} dx_j \tag{1-292}$$

This transformation is known as a ***contravariant*** transformation. However, not all quantities transform this way. For example the gradient is:

$$\nabla U = \frac{\partial U}{\partial x_1}\hat{\mathbf{i}} + \frac{\partial U}{\partial x_2}\hat{\mathbf{j}} + \frac{\partial U}{\partial x_3}\hat{\mathbf{k}} \tag{1-293}$$

whose transformation properties using the chain rule are:

$$\frac{\partial U'}{\partial x'_j} = \sum_i \frac{\partial U'}{\partial x_i}\frac{\partial x_i}{\partial x'_j} = \sum_i \frac{\partial U}{\partial x_i}\frac{\partial x_i}{\partial x'_j} \tag{1-294}$$

where we have used the fact that since U is a scalar $U = U'$. Note our transformation is "upside down" and in general can be written as:

$$dx_i = \sum_j \frac{\partial x_i}{\partial x'_j} dx'_j \tag{1-295}$$

This transformation is known as a ***covariant*** transformation.

We note that the difference between covariant and contravariant transformations does not create a problem in Cartesian coordinates because they are the same. We can see this from the transformation equations, Equations 1-260 through 1-262 and Equations 1-267 through 1-269, which indicate that:

$$\frac{\partial x'_i}{\partial x_j} = \frac{\partial x_j}{\partial x'_i} \tag{1-296}$$

and our transformations are the same. However, for other coordinate systems Equation 1-296 does not hold in general and the transformation differences between covariant and contravariant tensors become important. For example, if we transform from Cartesian Coordinates to polar coordinates in two dimensions we have:

$$x = r\cos\theta \tag{1-297}$$

$$y = r\sin\theta \tag{1-298}$$

The derivatives are:

$$dx = \cos\theta\, dr - r\sin\theta\, d\theta \tag{1-299}$$

$$dy = \sin\theta\, dr + r\cos\theta\, d\theta \tag{1-300}$$

which, we note we can write as:

$$dx = \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta \tag{1-301}$$

$$dy = \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \theta} d\theta \tag{1-302}$$

This is the form represented by Equation 1-294. To generalize, if a tensor transforms as:

$$A'_i = \sum_j \frac{\partial x_j}{\partial x'_i} A_j \tag{1-303}$$

or, more generally, as:

$$A'_{ijk\ldots} = \sum_l \sum_m \sum_n \cdots \frac{\partial x_l}{\partial x'_i}\frac{\partial x_m}{\partial x'_j}\frac{\partial x_n}{\partial x'_k} \cdots A_{lmn\ldots} \tag{1-304}$$

they are defined as covariant tensors. Tensors which transform as:

$$A'^{ijk\ldots} = \sum_l \sum_m \sum_n \cdots \frac{\partial x'_i}{\partial x_l}\frac{\partial x'_j}{\partial x_m}\frac{\partial x'_k}{\partial x_n}\cdots A^{lmn\ldots} \qquad (1\text{-}305)$$

are called contravariant tensors. It is also possible that the tensor is covariant in some indices and contravariant in others. Tensors of this type are called mixed tensors. By convention covariant tensors are written with indices as subscripts and contravariant tensors are written with indices as superscripts.

As we shall see, in physics applications there are several different kinds of tensors which arise from the theory, especially in the application of relativity to the electromagnetic field (Chapter XIII). One type of tensor is the ***symmetric tensor*** which has the property that the components have the relationship that $A_{ij} = A_{ji}$. Thus a second rank symmetric tensor, of 16 possible components, will have at most 10 different components. A symmetric tensor also has the property that it can be diagonalized by a suitable coordinate rotation. In this case the components of the diagonalized tensor are given by $A'_{ij} = A_i \delta_{ij}$. The other tensor which arises often is an ***antisymmetric tensor***. In this case $A_{ij} = -A_{ji}$ and there can be at most six different components for a second rank tensor. It is evident that the diagonal elements of an antisymmetric tensor must be zero.

Problem 1-15

Show that **1a** = **a1** if:

$$\mathbf{a} = \begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix}$$

Problem 1-16

Given two matrices:

$$\mathbf{a} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix};\ \mathbf{b} = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}$$

Find:

(a) $\mathbf{a}+\mathbf{b}$ (b) $\mathbf{a}-\mathbf{b}$ (c) $\mathbf{ab}$ (d) $\mathbf{ba}$ (e) $\tilde{\mathbf{a}}\mathbf{a}$ (f) $\tilde{\mathbf{b}}\mathbf{b}$ (g) $\mathbf{a}\tilde{\mathbf{a}}$ (h) $\mathbf{b}\tilde{\mathbf{b}}$

Problem 1-17

Given the two matrices:

$$\mathbf{a} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix};\ \mathbf{b} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 3 & -1 \end{bmatrix}$$

Find:

(a) $\mathbf{a}+\mathbf{b}$ (b) $\mathbf{a}-\mathbf{b}$ (c) $\mathbf{ab}$ (d) $\mathbf{ba}$ (e) $\tilde{\mathbf{a}}\mathbf{a}$ (f) $\tilde{\mathbf{b}}\mathbf{b}$ (g) $\mathbf{a}\tilde{\mathbf{a}}$ (h) $\mathbf{b}\tilde{\mathbf{b}}$

Problem 1-18

Show that the inner product, $\mathbf{A}\bullet\mathbb{T}$, is a vector.

Problem 1-19

Show that T'_{11}, found by expanding Equation 1-221 is the same as T'_{11} found from Equation 1-236.

Problem 1-20

Show that $\mathbf{R}\tilde{\mathbf{R}} = \mathbf{1}$.

1-13: Complex Numbers

The final topic we will cover in this mathematical introduction is that of complex numbers. Historically these numbers first came up when considering the solution of relatively simple quadratic equations. For example, the equation $x^2 - 1 = 0$ can be factored into $(x - 1)(x + 1) = 0$ and therefore has the two solutions $x = \pm 1$.

If we make a very simple change of sign so that the equation is now $x^2 + 1 = 0$ we cannot solve it using regular numbers. We could, however, define another kind of number. We call this number an imaginary number, which we represent by the italic symbol i and which satisfies the equation $i^2 = -1$. Then our equation becomes $x^2 - i^2 = 0$, which factors to $(x + i)(x - i) = 0$ and $x = \pm i$.

In the general case of the binomial equation the solutions to the general quadratic equation $ax^2 + bx + c = 0$ are:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{1-306}$$

Depending upon the nature of the values of a,b or c the solutions can be both real, both imaginary or, more generally, a combination of both. To include all possibilities, therefore, we define a ***complex number*** as:

$$z = x + iy \tag{1-307}$$

and with this addition to the set of real numbers a solution to the quadratic equation can always be determined.

Complex numbers, over and above their property of giving a fuller meaning to the solution of the quadratic equation, have other properties which make them very useful in the description of physical laws.

A complex number adds as an ordered pair of numbers:

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \tag{1-308}$$

It thus obeys the commutative and associative laws of addition.

Given two complex numbers, $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, they will be equal if and only if $x_1 = x_2$ and $y_1 = y_2$.

Multiplication of two complex numbers is defined as:

$$z_1 \times z_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2) \tag{1-309}$$

This comes directly from multiplying in the regular manner:

$$z_1 \times z_2 = (x_1 + iy_1) \times (x_2 + iy_2) \tag{1-310}$$

$$= x_1x_2 + iy_1x_2 + ix_1y_2 + i^2y_1y_2 \tag{1-311}$$

$$= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2) \tag{1-312}$$

as we had defined it. The complex numbers obey the commutative, associative, and distributive laws of multiplication. However, we note that multiplication mixes the real and imaginary components of the complex number, whereas addition and subtraction do not.

Division of two complex numbers is defined to be:

$$\frac{z_1}{z_2} = \frac{x_1x_2+y_1y_2}{x_2^2+y_2^2} + i\,\frac{(y_1x_2-x_1y_2)}{x_2^2+y_2^2} \qquad (1\text{-}313)$$

which can be derived by multiplying the top and the bottom of the term z_1/z_2 by $x_2 - iy_2$. Thus:

$$\frac{z_1}{z_2} = \frac{(x_1+iy_1)}{(x_2+iy_2)} = \frac{(x_1+iy_1)(x_2-iy_2)}{(x_2+iy_2)(x_2-iy_2)} \qquad (1\text{-}314)$$

which when multiplied out will give us Equation 1-313. Again we see that the real and imaginary components are mixed.

The complex number:

$$z^* = x - iy \qquad (1\text{-}315)$$

is called the ***complex conjugate*** of the complex number, z, and is formed by replacing i with $-i$ wherever it is found in a complex number.

Complex numbers can also be expressed in polar coordinates (r,ϕ) or in what is sometimes called polar form. In this form a complex number is represented as:

$$z = re^{i\phi} \qquad (1\text{-}316)$$

To see how this relates to our previous expression given by Equation 1-314 we expand the exponential function in a Taylor Series. Since e^x when expanded is:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}\,\cdot \qquad (1\text{-}317)$$

Substituting $i\phi$ for x our series becomes

$$e^{i\phi} = 1 + \frac{i\phi}{1!} + \frac{(i\phi)^2}{2!} + \frac{(i\phi)^3}{3!} + \frac{(i\phi)^4}{4!} + \cdots \frac{(i\phi)^n}{n!} + \cdots \qquad (1\text{-}318)$$

Rearranging we have:

$$e^{i\phi} = \left(1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \cdots\right) + i\left(\phi - \frac{\phi^3}{3!} + \frac{\phi^5}{5!} - \cdots\right) \qquad (1\text{-}319)$$

where we have used the fact that $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, etc. From this series expansion and from the Taylor Series for the sine and cosine we therefore see that Equation 1-316 can be written as:

$$e^{i\phi} = \cos\phi + i\sin\phi \qquad (1\text{-}320)$$

which is known as ***Euler's formula***.

Thus we rewrite equation 1-323 as:

$$z = r(\cos\phi + i\sin\phi) \qquad (1\text{-}321)$$

and from Equation 1-307 we see that,

$$x = r\cos\phi \qquad (1\text{-}322)$$

and,

$$y = r\sin\phi \qquad (1\text{-}323)$$

Conversely;

$$r = \sqrt{x^2 + y^2} \qquad (1\text{-}324)$$

and,

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \qquad (1\text{-}325)$$

We note that if we take the product of a complex number, z, and its complex conjugate, z^* we will have:

$$zz^* = (x + iy)(x - iy) \qquad (1\text{-}326)$$

$$= x^2 + y^2 = r^2 \tag{1-327}$$

Thus the magnitude of a complex number is given simply by the relationship:

$$r = \sqrt{zz^*} \tag{1-328}$$

Figure 1-25 shows how a complex number can be plotted and both forms (x,y) and (r,ϕ) are illustrated.

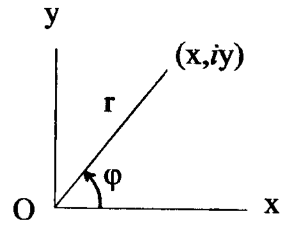


Figure 1-25: Graphical Representation of a Complex Number (the Argand Diagram)

We will come across complex numbers in the solution of differential equations, in the representation of wave motion and in other applications. Complex notation turns out to be a very convenient approach to representing some physical quantities. Often it is understood that only one part of the expression, either the real or the imaginary part, carries the physical (real) information. Sometimes the real part of the complex number represents one physical quantity and the imaginary part represents another. It is therefore important that the identity of the components be defined when using complex number notation.

If the mathematical operations are restricted to addition, subtraction, differentiation and integration the real and imaginary parts will not mix. Thus, quite often complicated expressions are simplified by complex notation. However, as Equations 1-312 and 1-314 indicate whenever complex numbers are multiplied or divided the real and imaginary components do mix. Thus, for this situation the complex notation will have to be converted to ordinary functions before a physically real result can be determined.

Problem 1-21

Find the equivalent complex number for the following quantities:

(a) $e^{\pi i}$ (b) $\cosh z$

(c) $\sin z$ (d) $\cos iz$

(e) r^{ik} (f) $e^{\frac{\pi}{2}i}$

Problem 1-22

Write in polar form:

(a) $z = 1 - i$ (b) $\ln z$

APPENDIX I

VECTOR IDENTITIES

$$a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + b\mathbf{B} \quad \text{(I-1)}$$

$$(a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A} \quad \text{(I-2)}$$

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \quad \text{(I-3)}$$

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \quad \text{(I-4)}$$

$$\mathbf{A}\bullet\mathbf{B} = \mathbf{B}\bullet\mathbf{A} \quad \text{(I-5)}$$

$$\mathbf{A}\bullet(\mathbf{B} + \mathbf{C}) = \mathbf{A}\bullet\mathbf{B} + \mathbf{A}\bullet\mathbf{C} \quad \text{(I-6)}$$

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A} \quad \text{(I-7)}$$

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C} \quad \text{(I-8)}$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A}\bullet\mathbf{C}) - \mathbf{C}(\mathbf{A}\bullet\mathbf{B}) \quad \text{(I-9)}$$

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{B}(\mathbf{A}\bullet\mathbf{C}) - \mathbf{A}(\mathbf{B}\bullet\mathbf{C}) \quad \text{(I-10)}$$

$$\mathbf{A}\bullet(\mathbf{B} \times \mathbf{C}) = \mathbf{B}\bullet(\mathbf{C} \times \mathbf{A}) = \mathbf{C}\bullet(\mathbf{A} \times \mathbf{B}) = -\mathbf{A}\bullet(\mathbf{C} \times \mathbf{B}) = -\mathbf{B}\bullet(\mathbf{A} \times \mathbf{C}) = -\mathbf{C}\bullet(\mathbf{B} \times \mathbf{A}) \quad \text{(I-11)}$$

$$\nabla(U + V) = \nabla U + \nabla V \quad \text{(I-12)}$$

$$\nabla(UV) = U\nabla V + V\nabla U \quad \text{(I-13)}$$

$$\nabla(\mathbf{A}\bullet\mathbf{B}) = \mathbf{B} \times (\nabla \times \mathbf{A}) + \mathbf{A} \times (\nabla \times \mathbf{B}) + (\mathbf{B}\bullet\nabla)\mathbf{A} + (\mathbf{A}\bullet\nabla)\mathbf{B} \quad \text{(I-14)}$$

$$\nabla\bullet(U\mathbf{A}) = \nabla U\bullet\mathbf{A} + U\nabla\bullet\mathbf{A} \quad \text{(I-15)}$$

$$\nabla\bullet(\mathbf{A} \times \mathbf{B}) = \mathbf{B}\bullet(\nabla \times \mathbf{A}) - \mathbf{A}\bullet(\nabla \times \mathbf{B}) \quad \text{(I-16)}$$

$$\nabla \times (\mathbf{A} + \mathbf{B}) = \nabla \times \mathbf{A} + \nabla \times \mathbf{B} \quad \text{(I-17)}$$

$$\nabla \times (U\mathbf{A}) = \nabla U \times \mathbf{A} + U(\nabla \times \mathbf{A}) \quad \text{(I-18)}$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B}\bullet\nabla)\mathbf{A} - (\mathbf{A}\bullet\nabla)\mathbf{B} + (\nabla\bullet\mathbf{B})\mathbf{A} - (\nabla\bullet\mathbf{A})\mathbf{B} \quad \text{(I-19)}$$

$$\nabla\bullet(\nabla \times \mathbf{A}) \equiv 0 \quad \text{(I-20)}$$

$$\nabla \times \nabla U \equiv 0 \quad \text{(I-21)}$$

$$\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla\bullet\mathbf{A}) - \nabla^2\mathbf{A} \quad \text{(I-22)}$$

$$(\mathbf{A}\bullet\nabla)\,\mathbf{r} = \mathbf{A} \quad \text{(I-23)}$$

$$\nabla(\mathbf{A}\bullet\mathbf{r}) = \mathbf{A} \quad \text{(I-24)}$$

$$\nabla\bullet\mathbf{r} = \nabla\bullet\mathbf{R} = 3 \quad \text{(I-25)}$$

$$\nabla \times \mathbf{r} = \nabla \times \mathbf{R} = 0 \quad \text{(I-26)}$$

$$\nabla\left(\frac{1}{R}\right) = -\frac{\hat{\mathbf{R}}}{R^2} = -\frac{\mathbf{R}}{R^3} \quad \text{(I-27)}$$

$$\nabla_{\mathbf{r}'}\left(\frac{1}{R}\right) = \frac{\hat{\mathbf{R}}}{R^2} = \frac{\mathbf{R}}{R^3} \quad \text{(I-28)}$$

$$\int_a^b (\nabla U)\bullet d\mathbf{l} = U(b) - U(a) \quad \text{(I-29)}$$

$$\oint_c \mathbf{A}\bullet d\mathbf{l} = \int_S (\nabla \times \mathbf{A})\bullet d\mathbf{S} \quad \text{(Stokes's Theorem)} \quad \text{(I-30)}$$

$$\oint_c U d\mathbf{l} = \int_S d\mathbf{S} \times \nabla U \quad \text{(I-31)}$$

$$\oint_c \mathbf{A} \times d\mathbf{l} = \int_S (\nabla\bullet\mathbf{A})d\mathbf{S} - \int_S \nabla(\mathbf{A}\bullet d\mathbf{S}) \quad \text{(I-32)}$$

$$\oint_C \mathbf{A}(\mathbf{B}\cdot d\mathbf{l}) = \int_S d\mathbf{S} \times (\nabla\cdot\mathbf{B})\mathbf{A} \qquad \text{(I-33)}$$

$$\int_\tau \nabla\cdot\mathbf{A}d\tau = \oint_S \mathbf{A}\cdot d\mathbf{S} \ \text{(Divergence Theorem)} \qquad \text{(I-34)}$$

$$\int_\tau \nabla U\, d\tau = \oint_S U\, d\mathbf{S} \qquad \text{(I-35)}$$

$$\int_\tau (\nabla\cdot\mathbf{B})\mathbf{A}d\tau = \oint_S \mathbf{A}(\mathbf{B}\cdot d\mathbf{S}) \qquad \text{(I-36)}$$

$$\int_\tau \nabla\cdot(U\mathbf{A})d\tau = \int_\tau [U(\nabla\cdot\mathbf{A}) + \mathbf{A}\cdot\nabla U]\, d\tau = \oint_S U(\mathbf{A}\cdot d\mathbf{S}) \qquad \text{(I-37)}$$

$$\int_\tau (\nabla \times \mathbf{A})d\tau = -\oint_S \mathbf{A} \times d\mathbf{S} = \oint_S d\mathbf{S} \times \mathbf{A} \qquad \text{(I-38)}$$

Problem AI-1
Verify the identity I-11
Problem AI-2
Verify the identity I-14
Problem AI-3
Find the expressions for each of the terms in the identity I-22 in Cartesian coordinates.

REFERENCES

Most texts on Electricity and Magnetism have a short introduction to the mathematics of the subject, especially vectors. The following texts cover many of the same topics we have:

Griffiths, D.J., Introduction to Electrodynamics, 3rd ed. (Prentice-Hall, Englewood, NJ, 1999)

Lorrain, P. and Corson, D., Electromagnetic Fields and Waves, 2nd ed. (Freeman, San Francisco, 1970)

Reitz, J.R., Milford, F.J. and Christy, R.W., Foundations of Electromagnetic Theory, 3rd ed. (Addison-Wesley, Reading, MA, 1979)

More general coverage of the topics in this chapter, including vectors, tensors and complex numbers, is found in books on mathematical physics such as:

Arfken, G., Mathematical Methods for Physicists, 3rd ed. (Academic Press, Orlando, Florida, 1985); Helmholtz's Theorem is discussed in Section 1.15.

Boas, M.L., Mathematical Methods in the Physical Sciences, 2nd ed. (Wiley, New York, 1983)

A book totally dedicated to vector calculus is:

Schey, H., Div, Grad, Curl and All That, 2nd ed. (Norton, New York, 1992)

Further discussion of the transformation properties of vectors and pseudovectors is found in the journal articles:

Quigley, R.J., Pseudovectors and Reflections, American Journal of Physics, **41**, 428-430 (1973)

Rosen, J., Transformation Properties of Electromagnetic Quantities under Space Inversion, Time Reversal, and Charge Conjugation, American Journal of Physics, **41**, 586-588 (1973)

Hauser, W., Vector Products and Pseudovectors, American Journal of Physics, **54**, 168-172 (1986)