Chapter 1. Vector Analysis

1.1. Introduction

This chapter will give an introduction to vector analysis. A **vector**, or displacement, has a **direction** and a **magnitude**. The following vector notations will be used in this course:

1. $\overline{A}$: the vector A

2. $\left|\overline{A}\right|$: the magnitude of vector A

The **opposite** of a vector $\overline{A}$ is a vector with the same magnitude as $\overline{A}$ but pointing in a direction opposite to that of $\overline{A}$. The opposite of vector $\overline{A}$ is written as $-\overline{A}$.

There are four different vector operations that we will be using in this course: vector addition, vector multiplication by a scalar, the dot product of two vectors, and the cross product of two vectors. We will start Chapter 1 by discussing these four vector operations in some detail.

1.1.1. Vector Addition

Two vectors $\overline{A}$ and $\overline{B}$ can be added. The result of this operation is a new vector $\overline{C}$ (see Figure 1.1a). Using vector notation we can write vector addition as follows:

$$\overline{A} + \overline{B} = \overline{C}$$

Vector addition is **commutative**. This means that the order in which two vectors are added does not affect the result (see Figure 1.1). The commutative properties of vector addition can be written as

$$\overline{A} + \overline{B} = \overline{C} = \overline{B} + \overline{A}$$

To subtract vector $\overline{B}$ from vector $\overline{A}$ is equivalent to adding the opposite of vector $\overline{B}$ to $\overline{A}$. In other words:

$$\overline{A} - \overline{B} = \overline{A} + \left(-\overline{B}\right)$$

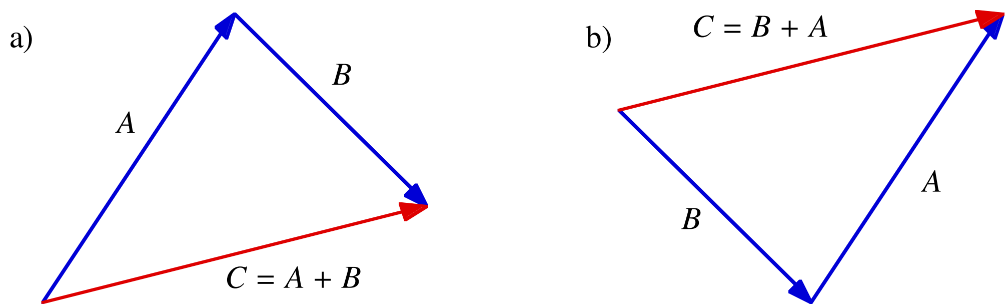

Figure 1.1. Vector Addition.

1.1.2. Multiplication by a scalar

A vector can be multiplied by a scalar. If the scalar a is a positive number (see Figure 1.2a) than the result of the multiplication of the vector $\overline{A}$ by a is a new vector with a magnitude equal to $a\left|\overline{A}\right|$ and a direction equal to the direction of $\overline{A}$. If the scalar a is a negative number (see Figure 1.2b) than the result of the multiplication of the vector $\overline{A}$ by a is a new vector with a magnitude equal to $|a| \cdot \left|\overline{A}\right|$ and a direction opposite to the direction of $\overline{A}$.

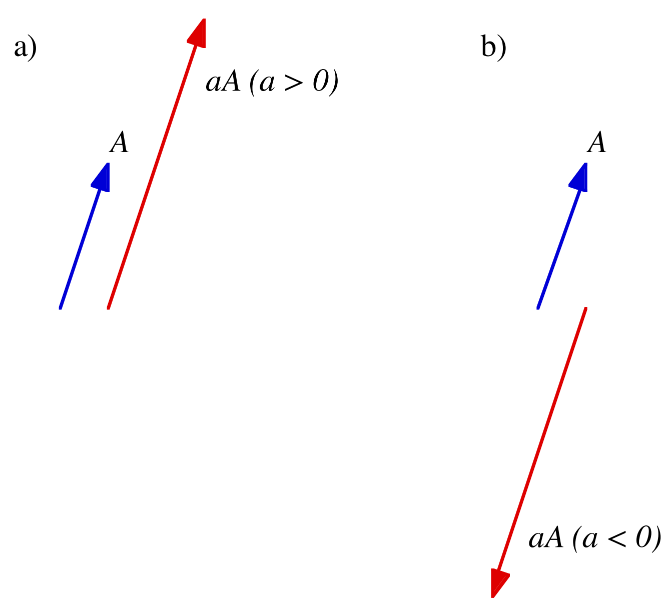

Figure 1.2. Vector multiplication.

Scalar multiplication is **distributive**. This means that the result of the multiplication of the vector sum of vector $\overline{A}$ and $\overline{B}$ by a scalar a is equal to the vector sum of $a\overline{A}$ and $a\overline{B}$:

$$a\left(\overline{A}+\overline{B}\right)=a\overline{A}+a\overline{B}$$

1.1.3. Dot product of two vectors

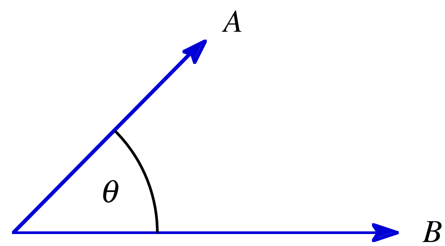

Figure 1.3. The scalar product of two vectors.

The dot product, also called the **scalar product**, is defined as

$$\overline{A}\bullet\overline{B}=\left|\overline{A}\right|\cdot\left|\overline{B}\right|\cdot\cos\theta$$

where θ is the angle between vectors $\overline{A}$ and $\overline{B}$ (see Figure 1.3). The dot product is **commutative** which means that the order of the vectors $\overline{A}$ and $\overline{B}$ does not effect the result of the dot product. In other words:

$$\overline{A}\bullet\overline{B}=\overline{B}\bullet\overline{A}$$

The dot product is also **distributive**:

$$\overline{A}\bullet\left(\overline{B}+\overline{C}\right)=\overline{A}\bullet\overline{B}+\overline{A}\bullet\overline{C}$$

The dot product will be frequently used to determine whether two vectors $\overline{A}$ and $\overline{B}$ are perpendicular. When $\overline{A}$ and $\overline{B}$ are perpendicular the angle θ is equal to 90°. The definition of the dot product shows that in this case the dot product between $\overline{A}$ and $\overline{B}$ is equal to zero.

1.1.4. Cross products of two vectors

The cross product of two vectors $\overline{A}$ and $\overline{B}$ is a third vector $\overline{C}$. The vector $\overline{C}$ is perpendicular to both $\overline{A}$ and $\overline{B}$, and has a length equal to

$$\left|\overline{C}\right|=\left|\overline{A}\times\overline{B}\right|=\left|\overline{A}\right|\cdot\left|\overline{B}\right|\cdot\sin\theta$$

where θ is the smallest angle between $\overline{A}$ and $\overline{B}$ (see Figure 1.4). The direction of the vector $\overline{C}$ can be determined using the **right-hand rule**. By applying the right-hand rule it can be shown easily that **the vector product is not commutative**. This implies that the order of the vectors $\overline{A}$ and $\overline{B}$ is important, and reversing the order will change the result of the vector product:

$$\overline{A} \times \overline{B} = -\overline{B} \times \overline{A}$$

The vector product is **distributive** which requires that

$$\overline{A} \times \left(\overline{B} + \overline{C}\right) = \overline{A} \times \overline{B} + \overline{A} \times \overline{C}$$

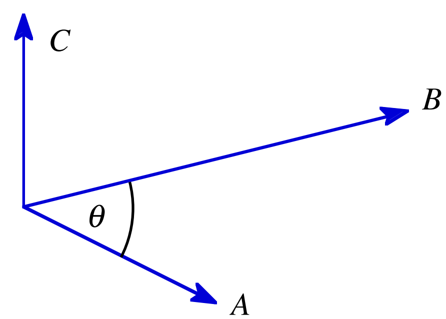

Figure 1.4. The vector product of $\overline{A}$ and $\overline{B}$.

Example: Problem 1.2

Is the cross product associative? If so, prove it; if not, provide a counter example.

If the cross product of two vectors is associative then the following relation must hold:

$$\left(\overline{A} \times \overline{B}\right) \times \overline{C} = \overline{A} \times \left(\overline{B} \times \overline{C}\right)$$

Consider the special case in which $\overline{A}$ = $\overline{B}$ and $\overline{C}$ is perpendicular to $\overline{A}$ and $\overline{B}$. In this case the cross product between $\overline{A}$ and $\overline{B}$ is equal to the null vector. The left-hand side of the equation is thus equal to the null vector. The cross product of $\overline{B}$ and $\overline{C}$ is a new vector, perpendicular to both $\overline{B}$ and $\overline{C}$, and with a length equal to $\left|\overline{B}\right| \cdot \left|\overline{C}\right|$. The cross product between $\overline{A}$ and $\overline{B} \times \overline{C}$ is a new vector. Since $\overline{A}$ and $\overline{B} \times \overline{C}$ are perpendicular, the magnitude of the cross product of $\overline{A}$ and $\overline{B} \times \overline{C}$ is equal to $\left|\overline{A}\right| \cdot \left|\overline{B}\right| \cdot \left|\overline{C}\right|$ which is non-zero. We thus conclude that $\left(\overline{A} \times \overline{B}\right) \times \overline{C}$ is not equal to $\overline{A} \times \left(\overline{B} \times \overline{C}\right)$ which shows that the cross product is not associative.

1.2. Vector components

A vector in three dimensions can be identified uniquely by specifying three coordinates. In a **Cartesian coordinate frame** three base vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$ are defined, parallel to the x, y, and z axes, respectively (see Figure 1.5). Each of these **base vectors** has unit length and is perpendicular to the other two base vectors. Any vector is uniquely defined by specifying its components along the x, y, and z axes. A vector $\overline{A}$ is defined in terms of the three Cartesian coordinates A_x, A_y, and A_z Using the three base vectors, one can reconstruct the vector $\overline{A}$:

$$\overline{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k}$$

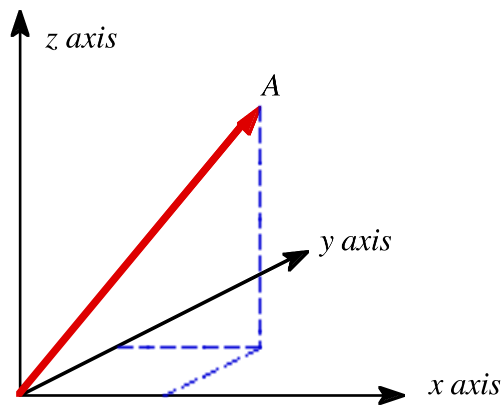

Figure 1.5. The Cartesian coordinate frame.

The vector operations discussed in Section 1.1 can be easily expressed in terms of vector components. We will now discuss each of these four vector operations.

1.2.1. Vector addition

The components of the vector sum of vectors $\overline{A}$ and $\overline{B}$ is equal to the sum of their like components. If a vector $\overline{C}$ is the vector sum of the vectors $\overline{A}$ and $\overline{B}$ then the components of $\overline{C}$ are equal to

$$C_x = A_x + B_x$$

$$C_y = A_y + B_y$$

$$C_z = A_z + B_z$$

In these equations, A_x, A_y, and A_z are the vector components of vector $\overline{A}$, and B_x, B_y, and B_z are the vector components of vector $\overline{B}$.

1.2.2. Multiplication by a scalar

The components of a vector $\overline{C}$, which is the result of a scalar multiplication of vector $\overline{A}$ with a scalar a, are equal to the components of $\overline{A}$ multiplied by a:

$$C_x = a \cdot A_x$$

$$C_y = a \cdot A_y$$

$$C_z = a \cdot A_z$$

1.2.3. Scalar product

The scalar product of vectors $\overline{A}$ and $\overline{B}$ is equal to the sum of the products of their like components:

$$\overline{A} \bullet \overline{B} = A_x B_x + A_y B_y + A_z B_z$$

This relation can be derived using the commutative properties of the scalar product:

$$\begin{aligned}
\overline{A} \bullet \overline{B} &= \left(A_x \hat{i} + A_y \hat{j} + A_z \hat{k}\right) \bullet \left(B_x \hat{i} + B_y \hat{j} + B_z \hat{k}\right) = \\
&= A_x B_x \left(\hat{i} \bullet \hat{i}\right) + A_y B_y \left(\hat{j} \bullet \hat{j}\right) + A_z B_z \left(\hat{k} \bullet \hat{k}\right) + \\
&\quad + \left(A_x B_y + A_y B_x\right)\left(\hat{i} \bullet \hat{j}\right) + \left(A_x B_z + A_z B_x\right)\left(\hat{i} \bullet \hat{k}\right) + \\
&\quad + \left(A_y B_z + A_z B_y\right)\left(\hat{j} \bullet \hat{k}\right) = \\
&= A_x B_x + A_y B_y + A_z B_z
\end{aligned}$$

In the last step of this derivation we have used the fact that the vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$ are perpendicular to each other and have unit length. The scalar products of these unit vectors are easy to evaluate and are equal to

$$\hat{i} \bullet \hat{i} = \hat{j} \bullet \hat{j} = \hat{k} \bullet \hat{k} = 1$$

$$\hat{i} \bullet \hat{j} = \hat{i} \bullet \hat{k} = \hat{j} \bullet \hat{k} = 0$$

1.2.4. Vector Product

The components of the vector $\overline{C}$, which is the vector product of vectors $\overline{A}$ and $\overline{B}$, can be calculated using the distributive properties of the vector product:

$$\begin{aligned}
\overline{A} \times \overline{B} &= \left(A_x\hat{i} + A_y\hat{j} + A_z\hat{k}\right) \times \left(B_x\hat{i} + B_y\hat{j} + B_z\hat{k}\right) = \\
&= A_xB_x\left(\hat{i} \times \hat{i}\right) + A_yB_y\left(\hat{j} \times \hat{j}\right) + A_zB_z\left(\hat{k} \times \hat{k}\right) + \\
&\quad + \left(A_xB_y - A_yB_x\right)\left(\hat{i} \times \hat{j}\right) + \left(A_xB_z - A_zB_x\right)\left(\hat{i} \times \hat{k}\right) + \\
&\quad + \left(A_yB_z - A_zB_y\right)\left(\hat{j} \times \hat{k}\right) = \\
&= \left(A_yB_z - A_zB_y\right)\hat{i} + \left(A_zB_x - A_xB_z\right)\hat{j} + \left(A_xB_y - A_yB_x\right)\hat{k}
\end{aligned}$$

In this derivation we have used the fact that the vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$ are perpendicular to each other and have unit length. The vector products of these unit vectors are easy to evaluate and are equal to

$$\hat{i} \times \hat{i} = \hat{j} \times \hat{j} = \hat{k} \times \hat{k} = 0$$

$$\hat{i} \times \hat{j} = -\hat{j} \times \hat{i} = \hat{k}$$

$$\hat{j} \times \hat{k} = -\hat{k} \times \hat{j} = \hat{i}$$

$$\hat{k} \times \hat{i} = -\hat{i} \times \hat{k} = \hat{j}$$

The expression of the vector product of $\overline{A}$ and $\overline{B}$ in terms of the components of these vectors can be neatly rewritten as a determinant:

$$\overline{A} \times \overline{B} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

Example: Problem 1.4

Use the vector product to find the components of the unit vector $\hat{n}$ perpendicular to the plane shown in Figure 1.6.

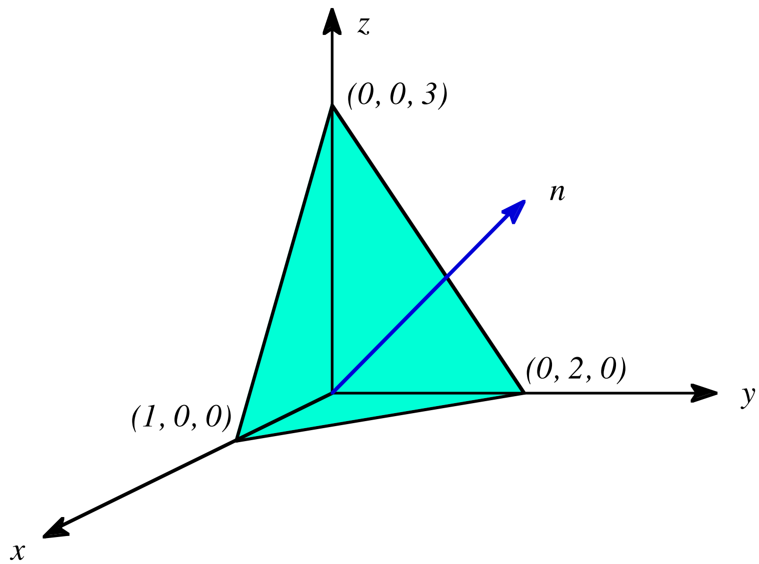

Figure 1.6. Problem 1.4

The orientation of a plane in three dimensions is completely determined by specifying two vectors that are parallel to this plane. Two possible vectors are the vector $\overline{A}$, connecting (1,0,0) and (0,2,0), and the vector $\overline{B}$, connecting (1,0,0) and (0,0,3). The vector product of $\overline{A}$ and $\overline{B}$ is a vector $\overline{C}$ which is perpendicular to both $\overline{A}$ and $\overline{B}$. The vector $\overline{C}$ will therefore be pointing in the same direction as $\hat{n}$. Using the definition of the vector product in terms of the components of vectors $\overline{A}$ and $\overline{B}$ we can calculate $\overline{C}$:

$$\overline{C} = \overline{A} \times \overline{B} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ -1 & 2 & 0 \\ -1 & 0 & 3 \end{vmatrix} = 6\hat{i} + 3\hat{j} + 2\hat{k}$$

The length of the vector $\overline{C}$ is equal to

$$\left|\overline{C}\right| = \sqrt{6^2 + 3^2 + 2^2} = 7$$

The unit vector $\hat{n}$ is parallel to $\overline{C}$ and has a length equal to 1. Therefore

$$\hat{n} = \frac{1}{7}\overline{C} = \frac{6}{7}\hat{i} + \frac{3}{7}\hat{j} + \frac{2}{7}\hat{k}$$

1.3. Vector Transformation

The three coordinates required to specify the direction and length of a vector are not uniquely defined. The same vector will in general have different coordinates in different coordinate systems. There is an infinite number of ways to choose a coordinate system, although the choice of the coordinate axes is usually influenced by the symmetry of the problem. Figure 1.7 illustrates two possible choices of the y and z axes. Coordinate system S is defined by the x, y, and z axes. Coordinate system S' is defined by the x', y', and z' axes. Coordinate system S' can be obtained by rotating coordinate system S through an angle ϕ around the x axis.

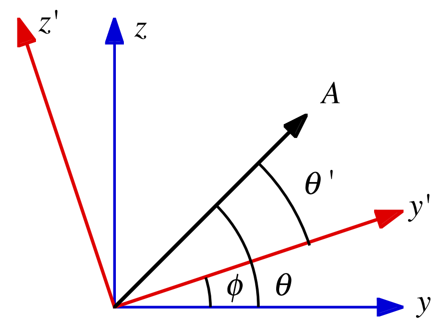

Figure 1.7. Coordinate Transformation.

The angle between the vector $\overline{A}$ and the y axis is equal to θ. The y and z coordinates of the vector $\overline{A}$ are therefore equal to

$$A_y = |A| \cdot \cos\theta$$

$$A_z = |A| \cdot \sin\theta$$

The angle between the vector $\overline{A}$ and the y' axis is equal to $\theta' = \theta - \phi$. The y' and z' coordinates of the vector $\overline{A}$ are therefore equal to

$$A_y' = |A| \cdot \cos(\theta') = |A| \cdot \cos(\theta - \phi)$$

$$A_z' = |A| \cdot \sin(\theta') = |A| \cdot \sin(\theta - \phi)$$

Using simple trigonometric relations we can obtain the component of the vector $\overline{A}$ in S' in terms of the components of the vector $\overline{A}$ in S:

$$A_y' = |A| \cdot \{\cos\theta \cos\phi + \sin\theta \sin\phi\} = A_y \cos\phi + A_z \sin\phi$$

$$A_z' = |A| \cdot \{-\cos\theta \sin\phi + \sin\theta \cos\phi\} = -A_y \sin\phi + A_z \cos\phi$$

These relations show, not unexpectedly, that the coordinates of the vector $\overline{A}$ in coordinate system S are related to the coordinates of vector $\overline{A}$ in coordinate system S'. This relation can be rewritten in matrix notation as

$$\begin{pmatrix} A_y' \\ A_z' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} A_y \\ A_z \end{pmatrix}$$

The rotation of the coordinate system S around the x axis leaves the x axis unchanged. The x coordinate of any vector in S will therefore be equal to the x' coordinate of this vector in S'. The three dimensional version of this coordinate transformation is therefore given by

$$\begin{pmatrix} A_x' \\ A_y' \\ A_z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}$$

The rotation of the coordinate system S around the x axis will not change the length of the vector $\overline{A}$. This can be verified by calculating the length of the vector $\overline{A}$ in coordinate system S. This length is equal to

$$\left|\overline{A}_S\right| = \sqrt{A_x^2 + A_y^2 + A_z^2}$$

The vector $\overline{A}$ in coordinate system S' is specified by the following coordinates:

$$\begin{pmatrix} A_x' \\ A_y' \\ A_z' \end{pmatrix} = \begin{pmatrix} A_x \\ A_y \cos\phi + A_z \sin\phi \\ -A_y \sin\phi + A_z \cos\phi \end{pmatrix}$$

The length of $\overline{A}$ in coordinate system S' is defined in terms of its coordinates in S' as

$$\left|\overline{A}_{S'}\right| = \sqrt{A_x'^2 + A_y'^2 + A_z'^2}$$

Using the relation between the coordinates of $\overline{A}$ in S' and the coordinates of $\overline{A}$ in S we obtain

$$\begin{aligned} \left|\overline{A}_{S'}\right| &= \sqrt{A_x^2 + \left(A_y \cos\phi + A_z \sin\phi\right)^2 + \left(-A_y \sin\phi + A_z \cos\phi\right)^2} = \\ &= \sqrt{A_x^2 + A_y^2 + A_z^2} = \left|\overline{A}_S\right| \end{aligned}$$

In general the coordinate transformation describing a rotation around an arbitrary axis can be written as

$$\begin{pmatrix} A_x' \\ A_y' \\ A_z' \end{pmatrix} = \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} R_{xx}A_x + R_{xy}A_y + R_{xz}A_z \\ R_{yx}A_x + R_{yy}A_y + R_{yz}A_z \\ R_{zx}A_x + R_{zy}A_y + R_{zz}A_z \end{pmatrix}$$

or, more compactly,

$$A_i' = \sum_{i=1}^{3} R_{ij} A_j$$

where $A_1 = A_x$, $A_2 = A_y$, and $A_3 = A_z$.

1.4. Differential calculus

The derivative df/dx of a function $f(x)$ tells us how rapidly the function varies when the argument x is changed by a tiny amount dx:

$$df = \left(\frac{df}{dx}\right) dx$$

A position dependent scalar function in three dimensions $f(x, y, z)$ will in general be a function of three variables. The variation of $f(x, y, z)$ between two closely spaced points depends not only on the distance between the two points, but also on their orientation. In this case, **the theorem of partial derivatives** can be used to calculate the change in $f(x, y, z)$:

$$df = \left(\frac{\partial f}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y}\right) dy + \left(\frac{\partial f}{\partial z}\right) dz$$

This equation can be rewritten as:

$$df = \left(\left(\frac{\partial f}{\partial x}\right)\hat{i} + \left(\frac{\partial f}{\partial y}\right)\hat{j} + \left(\frac{\partial f}{\partial z}\right)\hat{k}\right) \bullet \left((dx)\hat{i} + (dy)\hat{j} + (dz)\hat{k}\right)$$

The first term in this expression is called the **gradient** of f, written as $\overline{\nabla} f$, and is defined as

$$\overline{\nabla} f = \left(\frac{\partial f}{\partial x}\right)\hat{i} + \left(\frac{\partial f}{\partial y}\right)\hat{j} + \left(\frac{\partial f}{\partial z}\right)\hat{k}$$

The second term is called the **infinitesimal displacement vector** $d\bar{l}$ which is equal to

$$d\bar{l} = (dx)\hat{i} + \left(dy\right)\hat{j} + (dz)\hat{k}$$

The change in the scalar function *f* can thus be written as

$$df = \overline{\nabla} f \bullet d\bar{l}$$

From the definition of the gradient $\overline{\nabla} f$ it is clear that it is a vector. The **direction** of $\overline{\nabla} f$ points in the direction of maximum increase of the function *f*. This follows immediately from the expression of *df* in terms of $\overline{\nabla} f$:

$$df = \overline{\nabla} \bullet d\bar{l} = \left|\overline{\nabla} f\right| \cdot \left|d\bar{l}\right| \cdot \cos\theta$$

The change in *f*, *df*, will be maximum when $\theta = 0°$. In this case, $\overline{\nabla} f$ and $d\bar{l}$ are parallel. The **magnitude** of $\overline{\nabla} f$ gives the slope of the function *f* along the direction of maximum change. If the gradient of *f* vanishes at a certain point then the function *f* has a maximum, a minimum, a saddle point, or a shoulder at that point.

Example: Problem 1.13

Let $\bar{r}$ be the vector from some fixed point (x_0, y_0, z_0) to the point (x, y, z), and let r be its length. Show that

a) $\overline{\nabla}\left(r^2\right) = 2\bar{r}$

b) $\overline{\nabla}\left(\dfrac{1}{r}\right) = -\dfrac{\hat{r}}{r^2}$

c) What is he *general* formula for $\overline{\nabla}\left(r^n\right)$?

a) The vector $\bar{r}$ is equal to

$$\bar{r} = \left(x - x_0\right)\hat{i} + \left(y - y_0\right)\hat{j} + \left(z - z_0\right)\hat{k}$$

The scalar function r^2 is equal to the square of the length of the vector $\bar{r}$ and will be a function of x, y, and z:

$$r^2 = \left(x - x_0\right)^2 + \left(y - y_0\right)^2 + \left(z - z_0\right)^2$$

The gradient of r^2 can be obtained using the following partial derivatives:

$$\frac{\partial r^2}{\partial x} = 2(x - x_0)$$

$$\frac{\partial r^2}{\partial y} = 2(y - y_0)$$

$$\frac{\partial r^2}{\partial z} = 2(z - z_0)$$

The gradient of r^2 is equal to

$$\begin{aligned} \overline{\nabla} r^2 &= \frac{\partial r^2}{\partial x}\hat{i} + \frac{\partial r^2}{\partial y}\hat{j} + \frac{\partial r^2}{\partial z}\hat{z} = \\ &= 2(x - x_0)\hat{i} + 2(y - y_0)\hat{j} + 2(z - z_0)\hat{z} = 2r\hat{r} \end{aligned}$$

where $\hat{r}$ is the unit vector in the direction of $\bar{r}$.

b) The scalar function $1/r$ can be written as

$$\frac{1}{r} = \frac{1}{\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}}$$

The gradient of $1/r$ can be obtained using the following partial derivatives:

$$\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = \frac{-(x - x_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}}$$

$$\frac{\partial}{\partial y}\left(\frac{1}{r}\right) = \frac{-(y - y_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}}$$

$$\frac{\partial}{\partial z}\left(\frac{1}{r}\right) = \frac{-(z - z_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}}$$

The gradient of $1/r$ is thus equal to

$$\overline{\nabla}\left(\frac{1}{r}\right) = \frac{\partial}{\partial x}\left(\frac{1}{r}\right)\hat{i} + \frac{\partial}{\partial y}\left(\frac{1}{r}\right)\hat{j} + \frac{\partial}{\partial z}\left(\frac{1}{r}\right)\hat{k} =$$

$$= -\frac{(x-x_0)\hat{i} + (y-y_0)\hat{j} + (z-z_0)\hat{k}}{\left((x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2\right)^{3/2}} =$$

$$= -\frac{\overline{r}}{r^3} = -\frac{\hat{r}}{r^2}$$

c) The most general form of $\overline{\nabla}(r^n)$ can be guessed by comparing the results obtained in part a) and b). These answers suggest that the general form of $\overline{\nabla}(r^n)$ is given by

$$\overline{\nabla}(r^n) = nr^{n-1}\hat{r}$$

The operator $\overline{\nabla}$ has the formal appearance of a vector and can be written as

$$\overline{\nabla} = \hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}$$

The operator $\overline{\nabla}$ is also called the **vector operator**. If $\overline{\nabla}$ behaves like a vector than we expect that $\overline{\nabla}f$ behaves like a vector, and consequently rotates like a vector.

Example: Problem 1.14

Suppose that f is a function of two variables (y and z) only. Show that the gradient $\overline{\nabla}f$ transforms as a vector under a rotation about the x axis.

Consider a coordinate system S with x, y, and z axes. Coordinate system S is related to a coordinate system S' via a rotation by an angle ϕ about the x axis. The y and z coordinates of a vector in S are related to the y' and z' coordinates in S':

$$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} y\cos\phi + z\sin\phi \\ -y\sin\phi + z\cos\phi \end{pmatrix}$$

The gradient of f in the (y, z) frame is equal to

$$\overline{\nabla}f = \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}$$

The gradient of f in the (y', z') frame is equal to

$$\overline{\nabla} f' = \frac{\partial f}{\partial y'}\hat{j}' + \frac{\partial f}{\partial z'}\hat{k}'$$

If $\overline{\nabla} f$ transforms like a vector than the components of $\overline{\nabla} f'$ must be related to the components of $\overline{\nabla} f$ in the following manner:

$$\frac{\partial f}{\partial y'} = \frac{\partial f}{\partial y}\cos\phi + \frac{\partial f}{\partial z}\sin\phi$$

$$\frac{\partial f}{\partial z'} = -\frac{\partial f}{\partial y}\sin\phi + \frac{\partial f}{\partial z}\cos\phi$$

Using standard differential algebra we obtain the following relations for $\partial f/\partial y'$ and $\partial f/\partial z'$:

$$\frac{\partial f}{\partial y'} = \frac{\partial f}{\partial y}\frac{\partial y}{\partial y'} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y'}$$

$$\frac{\partial f}{\partial z'} = \frac{\partial f}{\partial y}\frac{\partial y}{\partial z'} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial z'}$$

To evaluate $\partial y/\partial y'$, $\partial y/\partial z'$, $\partial z/\partial y'$, and $\partial z/\partial z'$ we must express y and z in terms of y' and z'. This can be achieved by manipulating the following two relations:

$$y' = y\cos\phi + z\sin\phi$$

$$z' = -y\sin\phi + z\cos\phi$$

After a little algebra we obtain for y and z

$$y = y'\cos\phi - z'\sin\phi$$

$$z = y'\sin\phi + z'\cos\phi$$

Using these two relations we now can calculate the various partial derivatives of y and z:

$$\frac{\partial y}{\partial y'} = \cos\phi \qquad \frac{\partial y}{\partial z'} = -\sin\phi$$

$$\frac{\partial z}{\partial y'} = \sin\phi \qquad \frac{\partial z}{\partial z'} = \cos\phi$$

Using these partial derivatives we can now express $\overline{\nabla} f'$ in terms of $\overline{\nabla} f$:

$$\overline{\nabla} f_y' = \frac{\partial f}{\partial y'} = \cos\phi \frac{\partial f}{\partial y} + \sin\phi \frac{\partial f}{\partial z} = \cos\phi\left(\overline{\nabla} f_y\right) + \sin\left(\overline{\nabla} f_z\right)$$

$$\overline{\nabla} f_z' = \frac{\partial f}{\partial z'} = -\sin\phi \frac{\partial f}{\partial y} + \cos\phi \frac{\partial f}{\partial z} = -\sin\phi\left(\overline{\nabla} f_y\right) + \cos\left(\overline{\nabla} f_z\right)$$

These last two equations clearly show that $\overline{\nabla} f$ rotates like a vector.

If $\overline{\nabla}$ mimics a vector it may be used in vector multiplication. Three different types of vector multiplication can be carried out using $\overline{\nabla}$:

1. scalar multiplication with scalar function *f*, $\overline{\nabla} f$, also called the **gradient** of *f*.
2. dot (scalar) product with vector function $\overline{v}$, $\overline{\nabla} \bullet \overline{v}$, also called the **divergence** of $\overline{v}$.
3. cross (vector) product with vector function $\overline{v}$, $\overline{\nabla} \times \overline{v}$, also called the **curl** of $\overline{v}$.

We will now discuss in detail the divergence and curl of $\overline{v}$.

1.4.1. The Divergence

The divergence of a vector $\overline{v}$ is defined as

$$\begin{aligned} \overline{\nabla} \bullet \overline{v} &= \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) \bullet \left(v_x \hat{i} + v_y \hat{j} + v_z \hat{k}\right) = \\ &= \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \end{aligned}$$

and is a scalar. The geometrical interpretation of the divergence of a vector function $\overline{v}$ can be obtained by considering three special cases. First consider the vector function defined as

$$\overline{v}(x, y, z) = x\hat{i} + y\hat{j} + z\hat{k}$$

This vector function $\overline{v}$ is shown at a few points in the *x-y* plane in Figure 1.8a. The divergence of $\overline{v}$ is equal to

$$\overline{\nabla} \bullet \overline{v} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 1 + 1 + 1 = 3$$

Now consider a vector function $\bar{v}$ defined such that the vectors point toward the origin (see Figure 1.8b):

$$\bar{v}(x,y,z) = -x\hat{i} - y\hat{j} - z\hat{k}$$

For this vector function the divergence is equal to

$$\bar{\nabla} \bullet \bar{v} = \frac{\partial(-x)}{\partial x} + \frac{\partial(-y)}{\partial y} + \frac{\partial(-z)}{\partial z} = -1 - 1 - 1 = -3$$

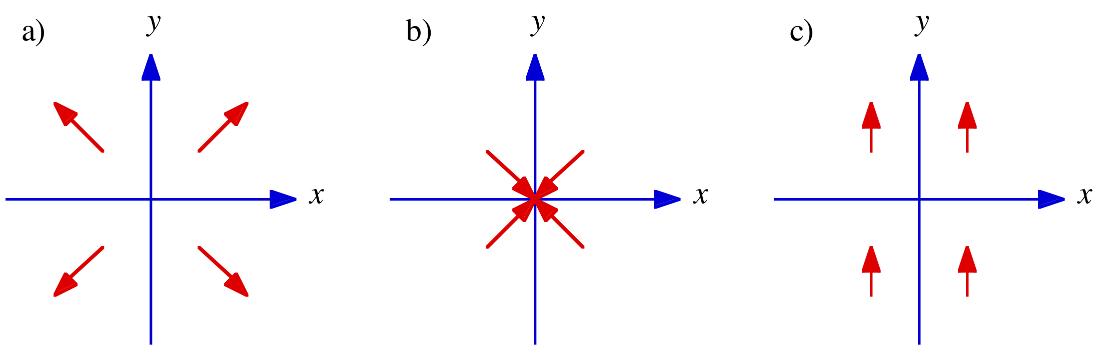

Figure 1.8. Various vector functions discussed in the text.

Finally consider the case in which the vector function $\bar{v}$ is a constant vector of unit length along the y axis (see Figure 1.8c), independent of position:

$$\bar{v}(x,y,z) = \hat{j}$$

For this vector function the divergence is equal to

$$\bar{\nabla} \bullet \bar{v} = \frac{\partial(0)}{\partial x} + \frac{\partial(1)}{\partial y} + \frac{\partial(0)}{\partial z} = 0$$

We conclude that the divergence of a vector function $\bar{v}$ evaluated at a particular point P is a measure of how much the vector function $\bar{v}$ spreads out:

1. $\bar{\nabla} \bullet \bar{v} > 0$: vector function $\bar{v}$ is spreading out.

2. $\bar{\nabla} \bullet \bar{v} < 0$: vector function $\bar{v}$ is narrowing in.

3. $\overline{\nabla} \bullet \overline{v} = 0$: vector function $\overline{v}$ is not spreading out or narrowing in.

One vector field that we will be using frequently is the electric field $\overline{E}$. The three cases just discussed are relevant in electrostatics:

1. the divergence of $\overline{E}$ generated by a positive point charge is positive (see Figure 1.9a).

2. the divergence of $\overline{E}$ generated by a negative point charge is negative (see Figure 1.9b).

3. the divergence of $\overline{E}$ generated by an infinitely large parallel-plate capacitor is zero (see Figure 1.9c).

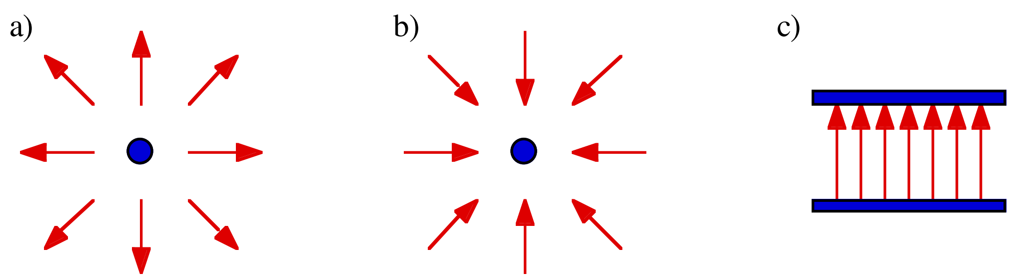

Figure 1.9. a) Electric field generated by a positive point charge. b) Electric field generated by a negative point charge. c) Electric field generated by an infinitely large parallel-plate capacitor.

Example: Problem 1.17

In two dimensions, show that the divergence transforms as a scalar under rotations. *Hint*: use the method of Problem 1.14 to calculate the derivatives.

If the divergence of a vector function $\overline{v}$ transforms as a scalar under rotation then the divergence of $\overline{v}$ in a coordinate system $S(x, y, z)$ must be identical to the divergence of $\overline{v}$ in a coordinate system $S'(x', y', z')$. This requires that

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = \frac{\partial v_x'}{\partial x'} + \frac{\partial v_y'}{\partial y'} + \frac{\partial v_z'}{\partial z'}$$

Consider a rotation about the x axis. In this case, $x = x'$ and $v_x = v_x'$, and consequently

$$\frac{\partial v_x}{\partial x} = \frac{\partial v_x'}{\partial x'}$$

The y' and z' components of vector $\bar{v}'$ are related to the y and z components of vector $\bar{v}$ in the following manner:

$$v_y' = v_y \cos\phi + v_z \sin\phi$$

$$v_z' = -v_y \sin\phi + v_z \cos\phi$$

The partial derivative of v_y' with respect to y' is equal to

$$\frac{\partial v_y'}{\partial y'} = \left(\frac{\partial v_y}{\partial y}\frac{\partial y}{\partial y'} + \frac{\partial v_y}{\partial z}\frac{\partial z}{\partial y'}\right)\cos\phi + \left(\frac{\partial v_z}{\partial y}\frac{\partial y}{\partial y'} + \frac{\partial v_z}{\partial z}\frac{\partial z}{\partial y'}\right)\sin\phi$$

Using the expressions for $\partial y/\partial y'$ and $\partial z/\partial y'$ obtained in problem 1.14 we can rewrite this expression as

$$\frac{\partial v_y'}{\partial y'} = \frac{\partial v_y}{\partial y}\cos^2\phi + \frac{\partial v_y}{\partial z}\cos\phi\sin\phi + \frac{\partial v_z}{\partial y}\cos\phi\sin\phi + \frac{\partial v_z}{\partial z}\sin^2\phi$$

In a similar manner the partial derivative of v_z' with respect to z' can be obtained:

$$\frac{\partial v_z'}{\partial z'} = \frac{\partial v_y}{\partial y}\sin^2\phi - \frac{\partial v_y}{\partial z}\cos\phi\sin\phi - \frac{\partial v_z}{\partial y}\cos\phi\sin\phi + \frac{\partial v_z}{\partial z}\cos^2\phi$$

Combining these last two equations, we obtain

$$\frac{\partial v_y'}{\partial y'} + \frac{\partial v_z'}{\partial z'} = \frac{\partial v_y}{\partial y}\left(\cos^2\phi + \sin^2\phi\right) + \frac{\partial v_z}{\partial z}\left(\cos^2\phi + \sin^2\phi\right) = \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$

and therefore, the divergence of a vector function $\bar{v}$ translates like a scalar under rotation.

1.4.2. The curl

The curl of vector function $\bar{v}$ is equal to

$$\bar{\nabla} \times \bar{v} = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right)\hat{i} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right)\hat{j} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right)\hat{k}$$

The curl of a vector function $\bar{v}$ evaluated at a certain point P is a measure of how much the vector function $\bar{v}$ curls around this point. For the vector functions $\bar{v}$ shown in Figure 1.8 the curl is zero.

Example: The magnetic field

Consider a wire of radius R carrying a current I along the z axis. The magnetic field produced by this wire depends only on the distance to the center of the wire. Its strength is equal to

$$B(r) = \frac{\mu_0 I}{2\pi r} \qquad r \geq R$$

$$B(r) = \frac{\mu_0 I}{2\pi r}\frac{r}{R^2} \qquad r \leq R$$

and its direction is tangent to the circle centered on the wire (see Figure 1.10).

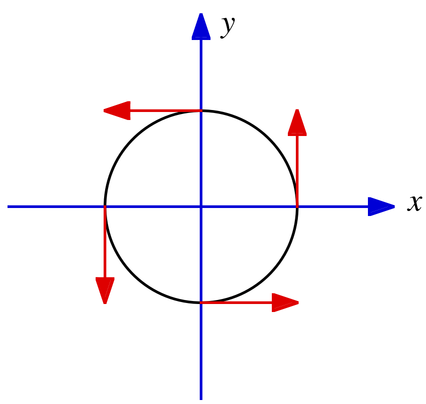

Figure 1.10. Magnetic vector field.

The magnetic field is a vector field, and the vector function is equal to

$$\bar{B}(x,y,z) = \frac{\mu_0 I}{2\pi\sqrt{x^2+y^2}}\left[\frac{y}{\sqrt{x^2+y^2}}\hat{i} - \frac{x}{\sqrt{x^2+y^2}}\hat{j}\right] \qquad x^2+y^2 \geq R$$

$$\bar{B}(x,y,z) = \frac{\mu_0 I}{2\pi R^2}\left[y\hat{i} - x\hat{j}\right] \qquad x^2+y^2 \leq R$$

In order to calculate the curl of $\bar{B}$ at $r > R$ we have to evaluate $\partial B_y/\partial x$ and $\partial B_x/\partial y$ at $r > R$:

$$\frac{\partial B_x}{\partial y} = \frac{\mu_0 I}{2\pi}\frac{\partial}{\partial y}\left[\frac{y}{x^2+y^2}\right] = \frac{\mu_0 I}{2\pi}\left[\frac{1}{x^2+y^2} - \frac{2y^2}{\left(x^2+y^2\right)^2}\right] = \frac{\mu_0 I}{2\pi}\frac{x^2-y^2}{\left(x^2+y^2\right)^2}$$

$$\frac{\partial B_y}{\partial x} = \frac{\mu_0 I}{2\pi}\frac{\partial}{\partial x}\left[\frac{-x}{x^2+y^2}\right] = \frac{\mu_0 I}{2\pi}\left[\frac{-1}{x^2+y^2} + \frac{2x^2}{\left(x^2+y^2\right)^2}\right] = \frac{\mu_0 I}{2\pi}\frac{x^2-y^2}{\left(x^2+y^2\right)^2}$$

The curl of $\overline{B}$ at $r > R$ is thus equal to

$$\overline{\nabla} \times \overline{B} = \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)\hat{k} = \frac{\mu_0 I}{2\pi}\left(\frac{x^2-y^2}{\left(x^2+y^2\right)^2} - \frac{x^2-y^2}{\left(x^2+y^2\right)^2}\right)\hat{k} = 0$$

In order to calculate the curl of $\overline{B}$ at $r < R$ we have to evaluate $\partial B_y/\partial x$ and $\partial B_x/\partial y$ at$r < R$:

$$\frac{\partial B_x}{\partial y} = \frac{\mu_0 I}{2\pi R^2}\frac{\partial}{\partial y}[y] = \frac{\mu_0 I}{2\pi R^2}$$

$$\frac{\partial B_y}{\partial x} = \frac{\mu_0 I}{2\pi R^2}\frac{\partial}{\partial x}[-x] = -\frac{\mu_0 I}{2\pi R^2}$$

The curl of $\overline{B}$ at $r < R$ is thus equal to

$$\overline{\nabla} \times \overline{B} = \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)\hat{k} = \left(\frac{\mu_0 I}{2\pi R^2} - \left(-\frac{\mu_0 I}{2\pi R^2}\right)\right)\hat{k} = \frac{\mu_0 I}{\pi R^2}\hat{k}$$

These calculations suggest that the curl of $\overline{B}$ evaluated at a certain point is proportional to the current density at that point: outside the wire ($r > R$) the current density is 0, while inside the wire ($r < R$) the current density is $I/(\pi R^2)$.

1.4.3. Product Rules

Without proof I mention the following product rules involving $\overline{\nabla}$ that will be used frequently in this course:

1. $\overline{\nabla}(fg) = f \cdot \overline{\nabla} g + g \cdot \overline{\nabla} f$: f and g are scalar functions

2. $\overline{\nabla}\left(\overline{A} \bullet \overline{B}\right) = \overline{A} \times \left(\overline{\nabla} \times \overline{B}\right) + \overline{B} \times \left(\overline{\nabla} \times \overline{A}\right) + \left(\overline{A} \bullet \overline{\nabla}\right)\overline{B} + \left(\overline{B} \bullet \overline{\nabla}\right)\overline{A}$: $\overline{A}$ and $\overline{B}$ are vector functions

3. $\overline{\nabla} \bullet \left(f\overline{A} \right) = f\left(\overline{\nabla} \bullet \overline{A} \right) + \overline{A} \bullet \overline{\nabla} f$: f is a scalar function and $\overline{A}$ is a vector function

4. $\overline{\nabla} \bullet \left(\overline{A} \times \overline{B} \right) = \overline{B} \bullet \left(\overline{\nabla} \times \overline{A} \right) - \overline{A} \bullet \left(\overline{\nabla} \times \overline{B} \right)$: $\overline{A}$ and $\overline{B}$ are vector functions

5. $\overline{\nabla} \times \left(f\overline{A} \right) = f\left(\overline{\nabla} \times \overline{A} \right) + \overline{A} \times \overline{\nabla} f$: f is a scalar function and $\overline{A}$ is a vector function

6. $\overline{\nabla} \times \left(\overline{A} \times \overline{B} \right) = \left(\overline{B} \bullet \overline{\nabla} \right)\overline{A} - \left(\overline{A} \bullet \overline{\nabla} \right)\overline{B} + \overline{A}\left(\overline{\nabla} \bullet \overline{B} \right) - \overline{B}\left(\overline{\nabla} \bullet \overline{A} \right)$: $\overline{A}$ and $\overline{B}$ are vector functions

1.4.4. Second derivatives

By applying $\overline{\nabla}$ twice we can construct five species of second derivatives:

1. The **divergence of a gradient**:

$$\overline{\nabla} \bullet \left(\overline{\nabla} T \right) = \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z} \right) \bullet \left(\hat{i}\frac{\partial T}{\partial x} + \hat{j}\frac{\partial T}{\partial y} + \hat{k}\frac{\partial T}{\partial z} \right) = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = \nabla^2 T$$

where T is a scalar function of x, y, and z. $\nabla^2 T$ is also called the **Laplacian** of T.

2. The **curl of a gradient** (is always zero):

$$\overline{\nabla} \times \left(\overline{\nabla} T \right) = 0$$

This can be shown easily:

$$\left[\overline{\nabla} \times \left(\overline{\nabla} T \right) \right]_x = \frac{\partial}{\partial y}\left(\overline{\nabla} T \right)_z - \frac{\partial}{\partial z}\left(\overline{\nabla} T \right)_y = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial z} T \right) - \frac{\partial}{\partial z}\left(\frac{\partial}{\partial y} T \right) = 0$$

3. The **gradient of a divergence**:

$$\overline{\nabla}\left(\overline{\nabla} \bullet \overline{v} \right)$$

4. The **divergence of a curl** (is always zero):

$$\overline{\nabla} \bullet \left(\overline{\nabla} \times \overline{v} \right) = 0$$

This can be shown easily:

$$\overline{\nabla} \bullet \left(\overline{\nabla} \times \overline{v}\right) = \frac{\partial}{\partial x}\left(\overline{\nabla} \times \overline{v}\right)_x + \frac{\partial}{\partial y}\left(\overline{\nabla} \times \overline{v}\right)_y + \frac{\partial}{\partial z}\left(\overline{\nabla} \times \overline{v}\right)_z =$$

$$= \frac{\partial}{\partial x}\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right) = 0$$

5. The **curl of a curl** of a vector function $\overline{v}$ can be expressed in terms of the Laplacian and the gradient of the divergence of $\overline{v}$:

$$\overline{\nabla} \times \left(\overline{\nabla} \times \overline{v}\right) = \overline{\nabla}\left(\overline{\nabla} \bullet \overline{v}\right) - \overline{\nabla}^2 \overline{v}$$

1.5. Integral calculus

Consider a one-dimensional function $df(x)/dx$ to be integrated between $x = a$ and $x = b$. This integral can be obtained using the **fundamental theorem of calculus**, which states that

$$\int_a^b \frac{df(x)}{dx} dx = f(b) - f(a)$$

In three dimensions this fundamental theorem is replaced by the **fundamental theorem for gradients**, which states that

$$\int_a^b \overline{\nabla} T \bullet d\overline{l} = T(b) - T(a)$$

In a one-dimensional world there is just a single path from a to b. In a three-dimensional world there are an infinite number of ways to move from a to b. Nevertheless, the **line integra**l of $\overline{\nabla} T$ depends only on the function values at the end points and not on the path chosen. Thus we conclude:

1. The line integral $\int_a^b \overline{\nabla} T \bullet d\overline{l}$ is independent of the path chosen between a and b.

2. The line integral $\oint \overline{\nabla} T \bullet d\overline{l}$ around any closed loop is zero.

Example 1.6

Although the line integral of the gradient of a vector function is independent of the path, the same is not true for arbitrary vector functions. Calculate

$$\int_a^b \bar{v} \bullet d\bar{l} \text{ from } a = (1, 1, 0) \text{ to } b = (2, 2, 0)$$

for the vector function $\bar{v} = y^2\hat{i} + 2x(y+1)\hat{j}$. Do it first by path (1), shown in Figure 1.11, and then by path (2).

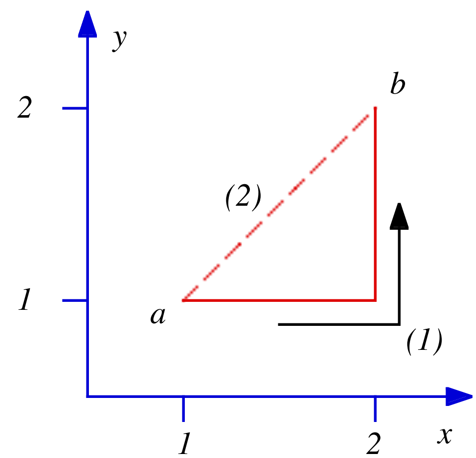

Figure 1.11. Example 1.6

Consider first path (1). The first part of this path is parallel to the x axis (y = 1) and along this segment the scalar product between $\bar{v}$ and $d\bar{l}$ is equal to

$$\bar{v} \bullet d\bar{l} = \left(y^2\hat{i} + 2x(y+1)\hat{j}\right) \bullet \left(dx\hat{i}\right) = y^2 dx = dx$$

The line integral along this segment is equal to

$$\int_{(1,1,0)}^{(2,1,0)} \bar{v} \bullet d\bar{l} = \int_1^2 dx = 2 - 1 = 1$$

The second part of path (1) is parallel to the y axis (x = 2) and along this segment the scalar product between $\bar{v}$ and $d\bar{l}$ is equal to

$$\bar{v} \bullet d\bar{l} = \left(y^2\hat{i} + 2x(y+1)\hat{j}\right) \bullet \left(dy\hat{j}\right) = 2x(y+1)dy = 4(y+1)dy$$

The line integral along this segment is equal to

$$\int_{(1,1,0)}^{(2,1,0)} \bar{v} \bullet d\bar{l} = \int_1^2 4(y+1)dy = \left(2y^2 + 4y\right)\Big|_1^2 = 16 - 6 = 10$$

The line integral of $\bar{v}$ along path (1) is equal to the sum of the line integrals along the two segments, and is thus equal to 1 + 10 = 11.

Consider now path (2). Along this path the x and y coordinates are equal ($x = y$ and $dx = dy$). The displacement vector $d\bar{l}$ along this path is equal to

$$d\bar{l} = dx\hat{i} + dx\hat{j}$$

The scalar product between $\bar{v}$ and $d\bar{l}$ along path (2) is equal to

$$\bar{v} \bullet d\bar{l} = \left(y^2\hat{i} + 2x(y+1)\hat{j}\right) \bullet \left(dx\hat{i} + dx\hat{j}\right) = \left(y^2 + 2x(y+1)\right)dx = \left(3x^2 + 2x\right)dx$$

The line integral along segment (2) is equal to

$$\int_{(1,1,0)}^{(2,1,0)} \bar{v} \bullet d\bar{l} = \int_1^2 \left(3x^2 + 2x\right)dx = \left(x^3 + x^2\right)\Big|_1^2 = 12 - 2 = 10$$

Comparing the line integral of $\bar{v}$ for path (1) and the line integral of $\bar{v}$ for path (2), we conclude that for the vector function $\bar{v}$ the line integral is path dependent.

In a three-dimensional world we can have, besides line integrals, surface integrals and volume integrals. Various fundamental theorems can simplify the calculation of surface and volume integrals. For example, the **fundamental theorem of divergences** states that the integral of a divergence of a vector function $\bar{v}$ over a volume is equal to the surface integral of the vector function $\bar{v}$ over the surface that bounds the volume:

$$\int_{Volume} \left(\bar{\nabla} \bullet \bar{v}\right) d\tau = \oint_{Surface} \bar{v} \bullet d\bar{a}$$

The right-hand side of this equation is called the **flux** of $\bar{v}$ through the surface that bounds the volume. The vector $d\bar{a}$ is a vector whose magnitude is equal to the area of an infinitesimal surface element and whose direction is perpendicular to the surface, pointing outwards. The left-hand side of this equation represents a **source term** (for example, $\bar{\nabla} \bullet \bar{E}$ integrated over a volume is proportional to the total charge present in that volume).

Example: Problem 1.32

Test the divergence theorem for the function $\bar{v} = xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}$. Take as your volume the cube shown in Figure 1.12, with sides of length 2.

The divergence of $\bar{v}$ is equal to

$$\overline{\nabla} \bullet \overline{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = y + 2z + 3x$$

The volume integral of $\overline{\nabla} \bullet \overline{v}$ over the volume of the cube is equal to

$$\int_{Volume} \left(\overline{\nabla} \bullet \overline{v}\right) d\tau = \int_0^2 dx \int_0^2 dy \int_0^2 (y + 2z + 3x) dz =$$

$$= \int_0^2 dx \int_0^2 (2y + 4 + 6x) dy =$$

$$= \int_0^2 (12 + 12x) dx = 48$$

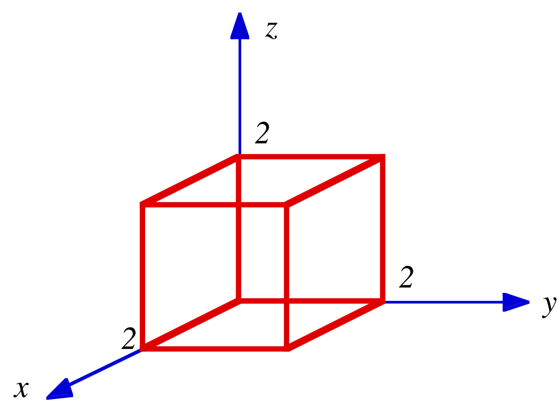

Figure 1.12. Problem 1.32.

To evaluate the surface integral of $\overline{v}$ across the surface of the cube, we have to consider each of the six faces separately. Start with considering the face of the cube in the *x-y* plane (with $z = 0$). The vector $d\overline{a}$ of this surface is equal to $d\overline{a} = -dxdy\hat{k}$. The scalar product of $\overline{v}$ and $d\overline{a}$ is equal to

$$\overline{v} \bullet d\overline{a} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet (-dxdy\hat{k}) = -3zxdxdy = 0$$

In the last step we have used the fact that $z = 0$ on this face. For the same reason the scalar product of $\overline{v}$ and $d\overline{a}$ is equal to zero on the face in the *x-z* plane and on the face in the *y-z* plane. On the face of the cube parallel to the *x-y* plane and with $z = 2$ the scalar product of $\overline{v}$ and $d\overline{a}$ is equal to

$$\overline{v} \bullet d\overline{a} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet (dxdy\hat{k}) = 3zxdxdy = 6xdxdy$$

The surface integral of $\bar{v}$ over this face is equal to

$$\int_{Surface} \bar{v} \bullet d\bar{a} = \int_0^2 dx \int_0^2 6xdy = \int_0^2 12xdx = 24$$

On the face the cube parallel to the x-z plane and with $y = 2$ the scalar product of $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet (dxdz\hat{j}) = 2yzdxdz = 4zdxdz$$

The surface integral of $\bar{v}$ over this face is equal to

$$\int_{Surface} \bar{v} \bullet d\bar{a} = \int_0^2 dx \int_0^2 4zdz = \int_0^2 8dx = 16$$

On the face of the cube parallel to the y-z plane and with $x = 2$ the scalar product of $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet (dydz\hat{i}) = xydydz = 2ydydz$$

The surface integral of $\bar{v}$ over this face is equal to

$$\int_{Surface} \bar{v} \bullet d\bar{a} = \int_0^2 dy \int_0^2 2ydz = \int_0^2 4ydy = 8$$

The surface integral across the surface of the cube is equal to the sum of the surface integrals across each of the 6 faces of the cube, and is thus equal to 24 + 16 + 8 = 48. This is equal to the volume integral of $\bar{\nabla} \bullet \bar{v}$ over the volume of the cube.

The **fundamental theorem for curls**, also known as Stokes' theorem, states that

$$\int_{Surface} \left(\bar{\nabla} \times \bar{v}\right) \bullet d\bar{a} = \oint_{Boundary} \bar{v} \bullet d\bar{l}$$

Here, $d\bar{a}$ is an infinitesimal surface element. It is a vector whose magnitude is equal to the area of the surface element and whose direction is perpendicular to the surface. The vector $d\bar{l}$ is tangential to the boundary of the surface. The orientation of the surface vector $d\bar{a}$ and the direction of integration of the boundary should be consistent with he right-hand rule. The following two corollaries follow from Stokes' theorem:

1. $\int_{Surface} \left(\overline{\nabla} \times \overline{v}\right) \bullet d\overline{a}$ depends only on the boundary line, not on the particular surface used.

2. $\oint_{Surface} \left(\overline{\nabla} \times \overline{v}\right) \bullet d\overline{a} = 0$ around any closed surface.

Example: Problem 1.33

Test Stokes' theorem for the vector function $\overline{v} = xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}$, using the triangular shaded area shown in Figure 1.13.

The curl of $\overline{v}$ is equal to

$$\overline{\nabla} \times \overline{v} = -2y\hat{i} - 3z\hat{j} - x\hat{k}$$

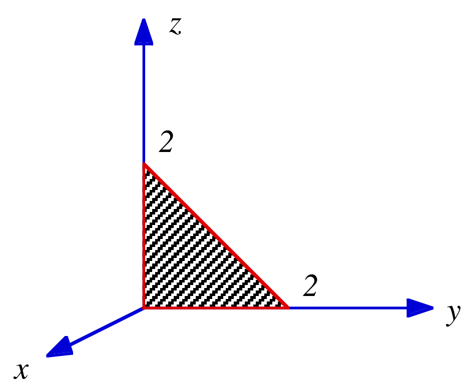

Fig. 1.13. Problem 1.33.

The surface vector $d\overline{a}$ is perpendicular to the surface and is equal to $d\overline{a} = dydz\hat{i}$. The scalar product between $\overline{\nabla} \times \overline{v}$ and $d\overline{a}$ is equal to

$$\left(\overline{\nabla} \times \overline{v}\right) \bullet d\overline{a} = \left(-2y\hat{i} - 3z\hat{j} - x\hat{k}\right) \bullet dydz\hat{i} = -2ydydz$$

The surface integral of $\overline{\nabla} \times \overline{v}$ is equal to

$$\int_{Surface} \left(\overline{\nabla} \times \overline{v}\right) \bullet d\overline{a} = \int_0^2 dy \int_0^{2-y} -2ydz = \int_0^2 2y\left(y-2\right)dy = -\frac{8}{3}$$

To evaluate the line integral around the boundary of the surface we have to evaluate the line integral along each of the three sides of the triangle. The direction of evaluation of the integral must be consistent with the chosen direction of $d\bar{a}$. The right-hand rule requires that the line integral is evaluated counter clockwise. First consider the segment between (0, 0, 0) and (0, 2, 0). Along this segment $d\bar{l} = dy\hat{j}$. The vector product between $\bar{v}$ and $d\bar{l}$ is equal to

$$\bar{v} \bullet d\bar{l} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet \left(dy\hat{j}\right) = 2yzdy = 0$$

since $z = 0$ along this segment. The line integral along this segment is therefore equal to zero. Now consider the line segment between (0, 2, 0) and (0, 0, 2). Along this segment $y + z = 2$ and $d\bar{l} = dy\hat{j} - dy\hat{k}$. The vector product between $\bar{v}$ and $d\bar{l}$ is equal to

$$\bar{v} \bullet d\bar{l} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet \left(dy\hat{j} - dy\hat{k}\right) = 2yzdy - 3zxdy = 2yzdy$$

since $x = 0$ along this segment. Along this segment $z = 2 - y$ and thus

$$\bar{v} \bullet d\bar{l} = 2y(2 - y)dy = \left(4y - 2y^2\right)dy$$

The line integral of $\bar{v}$ along this segment is thus equal to

$$\int_{Segment 2} \bar{v} \bullet d\bar{l} = \int_2^0 \left(4y - 2y^2\right)dy = \left(2y^2 - \frac{2}{3}y^3\right)\Big|_2^0 = -\frac{8}{3}$$

Note: the limits of the integral are chosen such that the line integral is evaluated in a counter clockwise direction. The third segment of the boundary to be considered connects (0, 0, 2) and (0, 0, 0). The vector $d\bar{l}$ along this segment is equal to $d\bar{l} = -dz\hat{k}$. The vector product between $\bar{v}$ and $d\bar{l}$ is equal to

$$\bar{v} \bullet d\bar{l} = \left(xy\hat{i} + 2yz\hat{j} + 3zx\hat{k}\right) \bullet \left(-dz\hat{k}\right) = -3zxdz = 0$$

since $x = 0$ along this segment. The line integral of $\bar{v}$ is therefore equal to zero. The line integral along the boundary of the surface is equal to the sum of the line integral of along each of the three segments. Thus

$$\int_{Boundary} \bar{v} \bullet d\bar{l} = 0 - \frac{8}{3} + 0 = -\frac{8}{3}$$

which is equal to the surface integral of $\bar{\nabla} \times \bar{v}$.

1.6. Curvilinear coordinates

The Cartesian coordinate system is a coordinate system that is often used in calculations involving systems with no apparent symmetry. To describe systems that have spherical or cylindrical symmetry it is often more convenient to use **spherical coordinates** or **cylindrical coordinates**, respectively. These two coordinate systems will be discussed in this section.

1.6.1. Spherical coordinates

Spherical coordinates are always used when the system under consideration has spherical symmetry. The location of a point P (see Figure 1.14) is completely determined by specifying the following three coordinates:

1. r : the distance between the origin and P.

2. θ : the **polar angle** which is the angle between the vector P and the z-axis (see Figure 1.14).

3. ϕ : the **azimuthal angle** which is the angle between the projection of the vector to P in the x-y plane and the x axis.

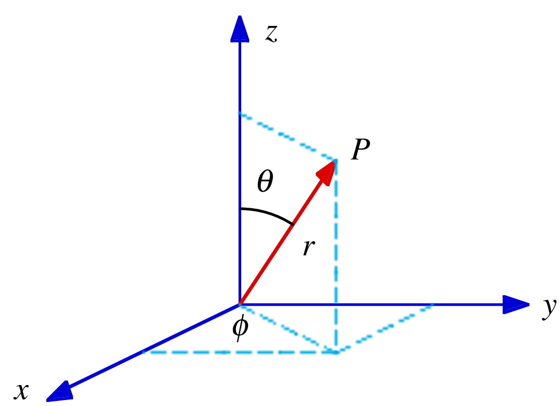

Figure 1.14. Spherical coordinates.

In general, any vector $\overline{A}$ can be expressed in terms of these three coordinates:

$$\overline{A} = A_r \hat{r} + A_\theta \hat{\theta} + A_\phi \hat{\phi}$$

In contrast to the unit vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$ in a Cartesian coordinate system, the unit vectors $\hat{r}$, $\hat{\theta}$, and $\hat{\phi}$ in a spherical coordinate system are not constant; they change direction as P moves around.

Consider a point P (see Figure 1.14) which is defined by the three spherical coordinates (r, θ, and ϕ). The corresponding Cartesian coordinates are

$$x = r\sin\theta\cos\phi$$

$$y = r\sin\theta\sin\phi$$

$$z = r\cos\theta$$

The unit vectors in the spherical coordinate system can be calculated as follows:

1. $\hat{r}$: The direction of this unit vector points from the origin to point P:

$$\hat{r} = \frac{1}{r}\left(x\hat{i} + y\hat{j} + z\hat{k}\right) = \left(\sin\theta\,\cos\phi\right)\hat{i} + \left(\sin\theta\,\sin\phi\right)\hat{j} + (\cos\theta)\hat{k}$$

2. $\hat{\theta}$: A change $d\theta$ in the polar angle θ will cause a change in the direction of the vector $\overline{P}$, without changing its length:

$$d\overline{P} = \left(r\cos\theta\cos\phi d\theta\right)\hat{i} + \left(r\cos\theta\sin\phi d\theta\right)\hat{j} - (r\sin\theta d\theta)\hat{k}$$

The unit vector $\hat{\theta}$ is defined as

$$\hat{\theta} = \frac{d\overline{P}}{\left|d\overline{P}\right|} = \left(\cos\theta\cos\phi\right)\hat{i} + \left(\cos\theta\sin\phi\right)\hat{j} - (\sin\theta)\hat{k}$$

3. $\hat{\phi}$: A change $d\phi$ in the polar angle ϕ will cause a change in the direction of the vector $\overline{P}$ without changing its length:

$$d\overline{P} = -\left(r\sin\theta\,\sin\phi d\phi\right)\hat{i} + \left(r\sin\theta\,\cos\phi d\phi\right)\hat{j}$$

The unit vector $\hat{\phi}$ is defined as

$$\hat{\phi} = \frac{d\overline{P}}{\left|d\overline{P}\right|} = -\left(\sin\phi\right)\hat{i} + \left(\cos\phi\right)\hat{j}$$

The expressions for $\hat{r}$, $\hat{\theta}$, and $\hat{\phi}$ clearly show that the direction of these unit vectors depends on the point being described. In contrast, in a Cartesian coordinate system the direction of the three unit vectors is fixed, and independent of the point being described.

An infinitesimal element of length in the $\hat{r}$ direction is simply dr:

$$dl_r = dr$$

The length of an infinitesimal element of length in the $\hat{\theta}$ direction (as a result of a change in the polar angle of $d\theta$) is equal to

$$dl_\theta = \sqrt{(r\cos\theta\cos\phi)^2 + (r\cos\theta\sin\phi)^2 + (-r\sin\theta)^2}\,d\theta = rd\theta$$

The length of an infinitesimal element of length in the $\hat{\phi}$ direction (as a result of a change in the azimuthal angle $d\phi$) is equal to

$$dl_\phi = \sqrt{(-r\sin\theta\sin\phi)^2 + (r\sin\theta\cos\phi)^2}\,d\phi = r\sin\theta d\phi$$

The most general infinitesimal displacement is thus equal to

$$d\bar{l} = dr \cdot \hat{r} + rd\theta \cdot \hat{\theta} + r\sin\theta d\phi \cdot \hat{\phi}$$

and this expression is used to evaluate line integrals in spherical coordinates. The infinitesimal volume element $d\tau$, in spherical coordinates, is the product of dl_r, dl_θ, and dl_ϕ:

$$d\tau = r^2 \sin\theta dr d\theta d\phi$$

There is no general expression for an infinitesimal surface element $d\bar{a}$ since it depends entirely on the orientation of the surface. For example, consider the surface of a sphere of radius r. On this surface r is constant, and the orientation of the surface is perpendicular to $\hat{r}$. In this case $d\bar{a}$ is equal to

$$d\bar{a} = dl_\theta dl_\phi \hat{r} = \left(r^2 \sin\theta d\theta d\phi\right)\hat{r}$$

Most calculations in spherical coordinates involve the application of vector derivatives. I will therefore summarize here, without derivation, the vector derivatives in spherical coordinates:

1. the **gradient** of a scalar function T:

$$\overline{\nabla} T = \frac{\partial T}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial T}{\partial \theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial T}{\partial \phi}\hat{\phi}$$

2. the **divergence** of a vector function $\overline{v}$:

$$\overline{\nabla} \bullet \overline{v} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 v_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta v_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\left(v_\phi\right)$$

3. the **curl** of a vector function $\overline{v}$:

$$\begin{aligned}\overline{\nabla} \times \overline{v} \quad = \quad & \frac{1}{r\sin\theta}\left(\frac{\partial}{\partial \theta}\left(\sin\theta v_\phi\right) - \frac{\partial}{\partial \phi}\left(v_\theta\right)\right)\hat{r} + \frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial \phi}\left(v_r\right) - \frac{\partial}{\partial r}\left(r v_\phi\right)\right)\hat{\theta} + \\ & + \frac{1}{r}\left(\frac{\partial}{\partial r}\left(r v_\theta\right) - \frac{\partial}{\partial \theta}\left(v_r\right)\right)\hat{\phi}\end{aligned}$$

4. the **Laplacian** of a scalar function *T*:

$$\overline{\nabla}^2 T = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial T}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 T}{\partial \phi^2}$$

1.6.2. Cylindrical coordinates

Cylindrical coordinates are always used when the system under consideration has cylindrical symmetry. The z axis is defined such that it coincides with the center axis of the cylinder. The location of a point *P* (see Figure 1.15) is completely determined by specifying the following three coordinates:

1. r : the distance between *P* and the z axis (see Figure 1.15).

2. ϕ : the azimuthal angle which is the angle between the projection of the vector to *P* in the *x-y* plane and the *x* axis (see Figure 1.15).

3. z : the z coordinate of point *P* (see Figure 1.15).

In general, any vector $\overline{A}$ can be expressed in terms of these three coordinates:

$$\overline{A} = A_r\hat{r} + A_\phi\hat{\phi} + A_z\hat{z}$$

However, the unit vectors $\hat{r}$ and $\hat{\phi}$ are not constant; they change direction as P moves around.

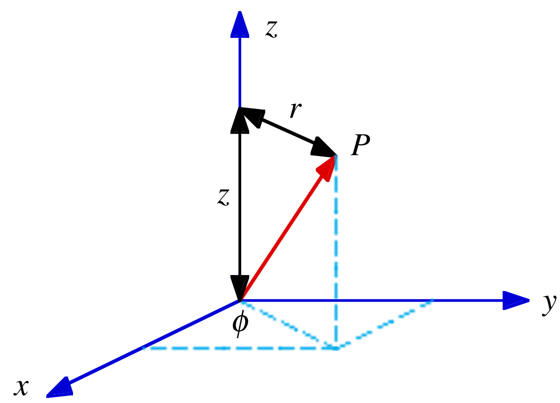

Figure 1.15. Cylindrical coordinates.

Consider a point P (see Figure 1.15) which is defined by three spherical coordinates (r, ϕ, and z). The corresponding Cartesian coordinates are

$$x = r\cos\phi$$

$$y = r\sin\phi$$

$$z = z$$

An infinitesimal change in r, ϕ, or z will produce the following infinitesimal displacements:

$$dl_r = dr$$

$$dl_\phi = \sqrt{(-r\sin\phi)^2 + (r\cos\phi)^2}\,d\phi = rd\phi$$

$$dl_z = dz$$

The infinitesimal displacement vector, to be used when evaluating line integrals in spherical coordinates, is equal to

$$d\bar{l} = dl_r\hat{r} + dl_\phi\hat{\phi} + dl_z\hat{z} = dr\hat{r} + rd\phi\hat{\phi} + dz\hat{z}$$

The infinitesimal volume element $d\tau$, to be used in volume integration, is equal to

$$d\tau = dl_r dl_\phi dl_z = rdrd\phi dz$$

The infinitesimal area element $d\bar{a}$, to be used in surface integration, is not uniquely determined and depends on the orientation of the surface. For example, if we carry out a surface integration over the surface of a cylinder of radius *r* (fixed), the infinitesimal surface element to be used is equal to

$$d\bar{a} = r\,d\phi\,dz\,\hat{r}$$

Most calculations in cylindrical coordinates involve the application of vector derivatives. I will therefore summarize here, without derivation, the vector derivatives in cylindrical coordinates:

1. The **gradient** of a scalar function *T*:

$$\overline{\nabla} T = \frac{\partial T}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial T}{\partial \phi}\hat{\phi} + \frac{\partial T}{\partial z}\hat{z}$$

2. The **divergence** of a vector function $\bar{v}$:

$$\overline{\nabla} \bullet \bar{v} = \frac{1}{r}\frac{\partial}{\partial r}(r v_r) + \frac{1}{r}\frac{\partial}{\partial \phi}(v_\phi) + \frac{\partial}{\partial z}(v_z)$$

3. The **curl** of a vector function $\bar{v}$:

$$\overline{\nabla} \times \bar{v} = \left(\frac{1}{r}\frac{\partial}{\partial \phi}(v_z) - \frac{\partial}{\partial z}(v_\phi)\right)\hat{r} + \left(\frac{\partial}{\partial z}(v_r) - \frac{\partial}{\partial r}(v_z)\right)\hat{\phi} + \frac{1}{r}\left(\frac{\partial}{\partial r}(r v_\phi) - \frac{\partial}{\partial \phi}(v_r)\right)\hat{z}$$

4. The **Laplacian** of a scalar function *T*:

$$\overline{\nabla}^2 T = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 T}{\partial \phi^2} + \frac{\partial^2 T}{\partial z^2}$$

Example: Problem 1.42

a) Find the divergence of the function

$$\bar{v} = r\left(2 + \sin^2\phi\right)\hat{r} + r\sin\phi\cos\phi\hat{\phi} + 3z\hat{z}$$

b) Test the divergence theorem for this function, using the quarter-cylinder (radius 2, height 5) shown in Figure 1.16.
c) Find the curl of $\bar{v}$.

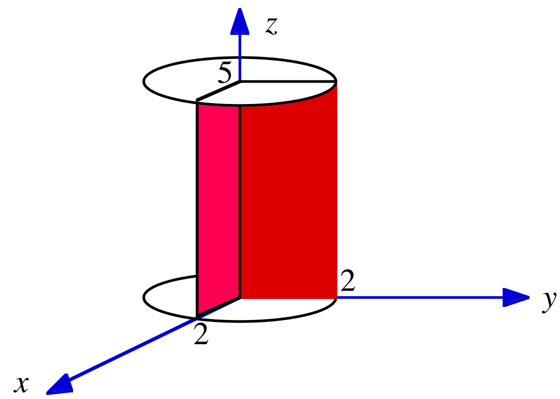

Figure 1.16. Problem 1.42.

a) To determine the divergence of $\bar{v}$ we have to calculate the following partial derivatives:

$$\frac{\partial}{\partial r}(rv_r) = \frac{\partial}{\partial r}\left(r^2\left(2+\sin^2\phi\right)\right) = 2r\left(2+\sin^2\phi\right)$$

$$\frac{\partial}{\partial \phi}\left(v_\phi\right) = \frac{\partial}{\partial \phi}\left(r\sin\phi\cos\phi\right) = r\left(1-2\sin^2\phi\right)$$

$$\frac{\partial}{\partial z}\left(v_z\right) = \frac{\partial}{\partial z}\left(3z^2\right) = 3$$

Using the definition of the divergence of $\bar{v}$ in terms of cylindrical coordinates we find that

$$\overline{\nabla} \bullet \bar{v} = 2\left(2+\sin^2\phi\right)+\left(1-2\sin^2\phi\right)+3 = 5+3 = 8$$

b) The volume integral of the divergence of $\bar{v}$ is equal to

$$\int\left(\overline{\nabla} \bullet \bar{v}\right)d\tau = \int_0^2 rdr\int_0^{\frac{\pi}{2}} d\phi\int_0^5 8dz = 40\int_0^2 dr\int_0^{\frac{\pi}{2}} d\phi = 40\cdot 2\cdot\frac{\pi}{2} = 40\pi$$

To calculate the surface integral of $\bar{v}$ we have to consider 5 different surfaces:

1. **The surface parallel to the *x-y* plane at *z* = 0.** The infinitesimal surface element for this surface is equal to

$$d\bar{a} = -dr\left(rd\phi\right)\hat{z}$$

(Note: the direction of $d\bar{a}$ is perpendicular to the surface and pointing outwards). The scalar product between $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = -3z dr(r d\phi) = 0$$

since $z = 0$ on this surface. The surface integral of $\bar{v}$ across this surface is therefore equal to zero.

2. **The surface parallel to the *x-y* plane at z = 5.** The infinitesimal surface element for this surface is equal to

$$d\bar{a} = dr(r d\phi)\hat{z}$$

The scalar product between $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = 3z dr(r d\phi) = 15 dr(r d\phi)$$

since $z = 5$ on this surface. The surface integral of $\bar{v}$ across this surface is equal to

$$\int \bar{v} \bullet d\bar{a} = \int_0^2 dr \int_0^{\frac{\pi}{2}} 15 r d\phi = \int_0^2 \frac{15\pi}{2} dr = 15\pi$$

3. **The surface in the *x-z* plane with ϕ = 0.** The infinitesimal surface element for this surface is equal to

$$d\bar{a} = -dr dz \hat{\phi}$$

The scalar product between $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = -r \sin\phi \cos\phi dr dz = 0$$

since $\phi = 0$ on this surface. The surface integral of $\bar{v}$ across this surface is therefore equal to zero.

4. **The surface in the *y-z* plane with ϕ = π/2.** The infinitesimal surface element for this surface is equal to

$$d\bar{a} = dr dz \hat{\phi}$$

The scalar product between $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = r\sin\phi\cos\phi drdz = 0$$

since $\phi = \pi/2$ on this surface. The surface integral of $\bar{v}$ across this surface is therefore equal to zero.

5. **The cylinder wall (r = 2).** The infinitesimal surface element for this surface is equal to

$$d\bar{a} = rd\phi dz\hat{r}$$

The scalar product between $\bar{v}$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = r\left(2 + \sin^2\phi\right)rd\phi dz = \left(8 + 4\sin^2\phi\right)d\phi dz$$

The surface integral of $\bar{v}$ across this surface is equal to

$$\begin{aligned}\int \bar{v} \bullet d\bar{a} &= \int_0^{\frac{\pi}{2}} d\phi \int_0^5 \left(8 + 4\sin^2\phi\right)dz = 5\int_0^{\frac{\pi}{2}} \left(8 + 4\sin^2\phi\right)d\phi = \\ &= 5\int_0^{\frac{\pi}{2}} (10 - 2\cos 2\phi)d\phi = 5(10 - \sin 2\phi)\Big|_0^{\frac{\pi}{2}} = 25\pi\end{aligned}$$

The surface integral of $\bar{v}$ across the whole surface is equal to the sum of the surface integral of $\bar{v}$ across each of these five surfaces:

$$\int \bar{v} \bullet d\bar{a} = 0 + 15\pi + 0 + 0 + 25\pi = 40\pi$$

which is equal to the volume integral of $\overline{\nabla} \bullet \bar{v}$.

1.7. The Dirac delta function

Consider the vector function $\bar{v}(r,\phi,\theta)$:

$$\bar{v}(r,\phi,\theta) = \frac{1}{r^2}\hat{r}$$

Consider the surface integral of $\bar{v}(r,\phi,\theta)$ across the surface of a sphere of radius R. The surface element vector $d\bar{a}$ for this surface is equal to

$$d\bar{a} = R^2 \sin\theta d\theta d\phi\hat{r}$$

The scalar product of $\bar{v}(r,\phi,\theta)$ and $d\bar{a}$ is equal to

$$\bar{v} \bullet d\bar{a} = \sin\theta d\theta d\phi$$

The surface integral of $\bar{v}(r,\phi,\theta)$ across the surface of this sphere is equal to

$$\int \bar{v} \bullet d\bar{a} = \int_0^{\pi} \sin\theta d\theta \int_0^{2\pi} d\phi = 2\pi \int_0^{\pi} \sin\theta d\theta = 4\pi$$

and is independent of R. Applying the divergence theorem we conclude that

$$\int_{Volume} \left(\overline{\nabla} \bullet \bar{v}\right) d\tau = 4\pi$$

for every sphere, centered on $r = 0$, and independent of the radius R. This suggests that the only contribution to the volume integral of $\overline{\nabla} \bullet \bar{v}$ comes from a single point at $r = 0$. The divergence of $\bar{v}$ is zero at every point except at $r = 0$:

$$\overline{\nabla} \bullet \bar{v} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{1}{r^2}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}(1) = 0$$

At $r = 0$, the divergence of $\bar{v}$ is undefined since 0/0 is undefined.

The function $\left(\overline{\nabla} \bullet \bar{v}\right)/4\pi$ is called the **Dirac delta function** $\delta(\bar{r})$ and will be used frequently in this course. The Dirac delta function has the following properties:

$$\int_{Volume} \delta(\bar{r}) d\tau = 1 \qquad \text{for any volume that includes } r = 0$$

$$\int_{Volume} \delta(\bar{r}) d\tau = 0 \qquad \text{for any volume that does not include } r = 0$$

The Dirac delta function $\delta(\bar{r})$ is thus defined as

$$\delta(\bar{r}) = \frac{1}{4\pi} \overline{\nabla} \bullet \left(\frac{1}{r^2}\hat{r}\right)$$

In Problem 1.13 we showed that the gradient of $1/r$ is equal to $-\hat{r}/r^2$. This relation can be used to rewrite the expression for the Dirac delta function as

$$\delta(r) = \frac{1}{4\pi} \overline{\nabla} \bullet \left(-\overline{\nabla}\left(\frac{1}{r}\right)\right) = -\frac{1}{4\pi} \overline{\nabla}^2 \left(\frac{1}{r}\right)$$

In this section we will discuss the use of the one- and three-dimensional Dirac delta functions.

1.7.1. The one-dimensional Dirac delta function

The one-dimensional Dirac delta function $\delta(x)$ is defined as

$$\delta(x) = 0 \quad if \quad x \neq 0$$

$$\delta(x) = \infty \quad if \quad x = 0$$

and its integral is equal to

$$\int_{-\infty}^{\infty} \delta(x)dx = 1$$

Any integral of $\delta(x)$ between $x = a$ and $x = b$ will be equal to 1 if $a < 0 < b$.

Every time we will be using the Dirac delta function it will be used as part of the integrand of an integral. For example:

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = \int_{-\infty}^{\infty} f(0)\delta(x)dx = f(0)\int_{-\infty}^{\infty} \delta(x)dx = f(0)$$

Example: Problem 1.43

Evaluate the following integrals:

a) $\int_2^6 \left(3x^2 - 2x - 1\right)\delta(x-3)dx$

b) $\int_0^5 \cos x \cdot \delta(x-\pi)dx$

c) $\int_0^3 x^3 \cdot \delta(x+1)dx$

d) $\int_{-\infty}^{\infty} \ln(x+3) \cdot \delta(x+2)dx$

a) The Dirac delta function $\delta(x-3)$ will be 0 for all x except $x = 3$. The limits of the integral include $x = 3$. Thus

$$\int_2^6 \left(3x^2 - 2x - 1\right)\delta(x-3)dx = 3(3)^2 - 2(3) - 1 = 20$$

b) The Dirac delta function $\delta(x-\pi)$ will be 0 for all x except $x = \pi$. The limits of the integral include $x = \pi$. Thus

$$\int_0^5 \cos x \cdot \delta(x-\pi)dx = \cos\pi = -1$$

c) The Dirac delta function $\delta(x+1)$ will be 0 for all x except x = - 1. The limits of the integral do not include x = - 1. Thus

$$\int_0^3 x^3 \cdot \delta(x+1)dx = 0$$

d) The Dirac delta function $\delta(x+2)$ will be 0 for all x except x = - 2. The limits of the integral include x = - 2. Thus

$$\int_{-\infty}^{\infty} \ln(x+3)\cdot\delta(x+2)dx = \ln(1) = 0$$

<u>Note</u>: Most mistakes made in the evaluation of integrals containing the Dirac delta function occur when the argument of the Dirac delta function has a form other than (x + a). This is best illustrated by looking at the following example:

$$\int_{-\infty}^{\infty} \delta(3x)dx = \int_{-\infty}^{\infty} \delta(3x)\frac{d(3x)}{3} = \frac{1}{3}\int_{-\infty}^{\infty} \delta(y)dy = \frac{1}{3}$$

1.7.2. The three-dimensional Dirac delta function

The three-dimensional Dirac delta function $\delta^3(\bar{r}) = \delta(x)\delta(y)\delta(z)$ is zero everywhere except at $\bar{r}$ = 0 (x = 0, y = 0, z = 0). Every time we will be using the three-dimensional Dirac delta function it will be used as part of the integrand of an integral. For example:

$$\int_V f(\bar{r})\delta^3(\bar{r})d\tau = f(0)$$

if the integration volume V includes $\bar{r}$ = 0.

Example: Problem 1.47a

Evaluate the following integral:

$$\int_V \left(r^2 + \bar{r} \bullet \bar{r}_0 + r_0{}^2\right)\delta^3\left(\bar{r} - \bar{r}_0\right)d\tau$$

The Dirac delta function is zero everywhere except at $\bar{r}$ = $\bar{r}_0$. The volume integral is thus equal to

$$\int_V \left(r^2 + \bar{r} \bullet \bar{r}_0 + r_0{}^2\right)\delta^3\left(\bar{r} - \bar{r}_0\right)d\tau = r_0{}^2 + \bar{r}_0 \bullet \bar{r}_0 + r_0{}^2 = 3r_0{}^2$$

1.8. Helmholtz Theorem

Consider being told that the divergence of a vector function $\overline{F}$ is the scalar function D and that the curl of the vector function $\overline{F}$ is the vector function $\overline{C}$. Is this sufficient information to determine the vector function $\overline{F}$ uniquely?

The vector function $\overline{F}$ satisfies the following relations:

$$\overline{\nabla} \bullet \overline{F} = D$$

$$\overline{\nabla} \times \overline{F} = \overline{C}$$

The divergence of $\overline{C}$ must be zero since the divergence of the curl of any vector function is zero ($\overline{\nabla} \bullet \left(\overline{\nabla} \times \overline{F}\right) = 0 = \overline{\nabla} \bullet \overline{C}$). Consider the following solution:

$$\overline{F} = -\overline{\nabla} U + \overline{\nabla} \times \overline{W}$$

where

$$U(\bar{r}) = \frac{1}{4\pi}\int \frac{D(\bar{r}\,')}{r - r'}d\tau'$$

$$\overline{W}(\bar{r}) = \frac{1}{4\pi}\int \frac{\overline{C}(\bar{r}\,')}{r - r'}d\tau'$$

If $\overline{F}$ is a correct solution than the divergence of $\overline{F}$ must be equal to D:

$$\begin{aligned}\overline{\nabla} \bullet \overline{F} \quad &= \quad -\overline{\nabla} \bullet \overline{\nabla} U + \overline{\nabla} \bullet \left(\overline{\nabla} \times \overline{W}\right) = -\overline{\nabla}^2 U = \\ &= \quad -\frac{1}{4\pi}\int D(r')\overline{\nabla}^2\left(\frac{1}{r - r'}\right)d\tau' = \int D(r')\delta^3(r - r')d\tau' = D(r)\end{aligned}$$

Here we have used one of the properties of second derivatives, which states that $\overline{\nabla} \bullet \left(\overline{\nabla} \times \overline{W}\right) = 0$ for any vector function $\overline{W}$.

If $\overline{F}$ is a correct solution than the curl of $\overline{F}$ must be equal to $\overline{C}$:

$$\overline{\nabla} \times \overline{F} = -\overline{\nabla} \times \overline{\nabla} U + \overline{\nabla} \times \left(\overline{\nabla} \times \overline{W}\right) = \overline{\nabla} \times \left(\overline{\nabla} \times \overline{W}\right) = -\overline{\nabla}^2 \overline{W}^2 + \overline{\nabla}\left(\overline{\nabla} \bullet \overline{W}\right)$$

Here we have used one of the properties of second derivatives which states that $\overline{\nabla} \times \overline{\nabla} U = 0$ for any scalar function *U*. The Laplacian of $\overline{W}$ is equal to

$$\overline{\nabla}^2 \overline{W}(\overline{r}) = \frac{1}{4\pi} \int \overline{C}(\overline{r}') \overline{\nabla}^2 \left(\frac{1}{r - r'} \right) d\tau' = -\int \overline{C}(\overline{r}') \delta^3(\overline{r} - \overline{r}') d\tau' = -\overline{C}(\overline{r})$$

The divergence of $\overline{W}$ is equal to

$$\begin{aligned} \overline{\nabla} \bullet \overline{W}(\overline{r}) &= \frac{1}{4\pi} \int_V \overline{C}(\overline{r}') \bullet \overline{\nabla} \left(\frac{1}{r - r'} \right) d\tau' = -\frac{1}{4\pi} \int_V \overline{C}(\overline{r}') \bullet \overline{\nabla}' \left(\frac{1}{r - r'} \right) d\tau' = \\ &= -\frac{1}{4\pi} \int_V \overline{\nabla}' \bullet \left(\frac{1}{r - r'} \overline{C}(\overline{r}') \right) d\tau' + \frac{1}{4\pi} \int_V \frac{1}{r - r'} \left(\overline{\nabla}' \bullet \overline{C}(\overline{r}') \right) d\tau' = \\ &= -\frac{1}{4\pi} \int_S \frac{1}{r - r'} \overline{C}(\overline{r}') \bullet d\overline{a}' \end{aligned}$$

In this derivation we have used the product rule for divergences (Griffiths page 21), the divergence theorem and the fact that the divergence of $\overline{C}$ is equal to zero. The surface integral of $\overline{C}/(r - r')$ will be equal to zero if $\overline{C}$ goes to zero faster than $1/r^2$ when $r \to \infty$. If this is the case, the divergence of $\overline{W}$ is equal to zero, and consequently

$$\overline{\nabla} \times \overline{F} = -\overline{\nabla}^2 \overline{W} + \overline{\nabla} \left(\overline{\nabla} \bullet \overline{W} \right) = \overline{C} + 0 = \overline{C}$$

So far we have verified that we can obtain a vector function $\overline{F}$ if its divergence and curl are given. However, the vector function $\overline{F}$ found is not unique. Consider a vector function $\overline{A}$ whose curl and divergence vanish. Since the curl is distributive we can express the curl of the new vector function in terms of the curl of $\overline{F}$ and the curl of $\overline{A}$:

$$\overline{\nabla} \times \left(\overline{F} + \overline{A} \right) = \overline{\nabla} \times \overline{F} + \overline{\nabla} \times \overline{A} = \overline{\nabla} \times \overline{F} = \overline{C}$$

Since the divergence is distributive we can express the divergence of the new vector function in terms of the divergence of $\overline{F}$ and the divergence of $\overline{A}$:

$$\overline{\nabla} \bullet \left(\overline{F} + \overline{A} \right) = \overline{\nabla} \bullet \overline{F} + \overline{\nabla} \bullet \overline{A} = \overline{\nabla} \bullet \overline{F} = D$$

However, there is no function that has zero divergence and zero curl and goes to zero at infinity. Since it is expected that the electric and magnetic fields go to zero at large distances we will exclude this type of contributions to $\overline{F}$ by requiring that $\overline{F}$ goes to zero at large distances. With this requirement, $\overline{F}$ is uniquely defined if its curl and divergence are given. This conclusion is know as the **Helmholtz theorem**:

If the divergence $D(\bar{r})$ and the curl $C(\bar{r})$ of a vector function $\overline{F}(\bar{r})$ are specified, and if they both go to zero faster than $1/r^2$ as $r \to \infty$ and if $\overline{F}(\bar{r})$ goes to zero as $r \to \infty$ then $\overline{F}(\bar{r})$ is given uniquely by

$$\overline{F} = -\overline{\nabla} U + \overline{\nabla} \times \overline{W}$$

where

$$U(\bar{r}) = \frac{1}{4\pi} \int \frac{D(\bar{r}')}{r - r'} d\tau'$$

$$\overline{W}(\bar{r}) = \frac{1}{4\pi} \int \frac{\overline{C}(\bar{r}')}{r - r'} d\tau'$$

1.9. Scalar and vector potentials

Consider a vector function $\overline{F}(\bar{r})$ which is the gradient of a scalar function $U(\bar{r})$: $\overline{F}(\bar{r}) = -\overline{\nabla} U(\bar{r})$. The scalar function $U(\bar{r})$ is also called the **scalar potential** of the field $\overline{F}(\bar{r})$. The vector field $\overline{F}(\bar{r})$, defined by the scalar function $U(\bar{r})$, has the following properties:

1. $\oint \overline{F}(\bar{r}) \bullet d\bar{l} = 0$ for any closed loop.

2. $\int_a^b \overline{F}(\bar{r}) \bullet d\bar{l}$ is independent of path, for any given set of end points.

3. $\overline{\nabla} \times \overline{F} = 0$ everywhere.

The vector field $\overline{F}(\bar{r})$ generated by the scalar function $U(\bar{r})$ is called a **curl-less field**. In Example 1.6 we showed that for the vector function $\bar{v} = y^2\hat{i} + 2x(y+1)\hat{j}$ the line integral is path dependent. The curl of this vector function is equal to $\overline{\nabla} \times \bar{v} = 2(y+1)\hat{i} - 2y\hat{k}$ and consequently this vector function can not be written as the gradient of a scalar function U. Therefore, the line integral of this vector function is path dependent, as was demonstrated in Example 1.6.

Now consider a vector field $\overline{F}(\bar{r})$ which is the curl of a vector function $\overline{W}$: $\overline{F} = \overline{\nabla} \times \overline{W}$. The vector function $\overline{W}$ is called the **vector potential** for the field $\overline{F}(\bar{r})$. The vector field $\overline{F}(\bar{r})$ defined by the vector potential W has the following properties:

1. $\oint_S \overline{F} \bullet d\overline{a} = 0$ for any closed surface.

2. $\int_S \overline{F} \bullet d\overline{a}$ is independent of the surface chosen, and depends only on the boundary line.

3. $\overline{\nabla} \bullet \overline{F} = 0$ everywhere.

The vector field $\overline{F}(\overline{r})$ generated by the vector potential $\overline{W}$ is called a **divergence-less field**.

The scalar potential U and the vector potential $\overline{W}$ are not unique. Any constant can be added to U without effecting its gradient, since the gradient of a constant is equal to zero. Any gradient of a scalar function can be added to $\overline{W}$ without effecting its curl, since the curl of a gradient is equal to zero.