## 1. VECTORS AND FIELDS

A *field* is a physical quantity that has a value associated with each point in space and time. The space and time-dependent values associated with scalar fields (such as temperature or density) are simply numbers, while those for vector fields (such as velocity, force, current density, electric field, or magnetic field) have both a scalar magnitude and a direction. Typeset vector fields will be denoted as boldfaced symbols (e.g., electric field $\mathbf{E}$), but sometimes they are denoted with an arrow above the symbol (e.g., $\overrightarrow{E}$).

Since vectors can point in any direction, we usually express them in component form by specifying their magnitude in three mutually orthogonal directions. Cartesian coordinates, having a basis set of unit vectors (dimensionless vectors of length one) pointing along the $x$, $y$, and $z$ axes are commonly used. Denoting the unit vector directions with hats (e.g., $\hat{x}$), an arbitrary vector is expressed as

$$\mathbf{A} = A_x\hat{x} + A_y\hat{y} + A_z\hat{z} = (A_x, A_y, A_z)$$

where $A_x$, $A_y$, and $A_z$ are the scalar magnitudes of the components of $\mathbf{A}$ along the positive $\hat{x}$, $\hat{y}$, and $\hat{z}$ directions, respectively. Note that for an arbitrary vector field, the magnitude of each vector component can itself be a function of location and time, such that $\mathbf{A}$ can be expressed most generally as:

$$\mathbf{A} = A_x(x,y,z,t)\hat{x} + A_y(x,y,z,t)\hat{y} + A_z(x,y,z,t)\hat{z}$$

The magnitude of vector $\mathbf{A}$ is denoted $|\mathbf{A}|$ (or sometimes just $A$), and is given by

$$A = |\mathbf{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2}$$

Sometimes it is convenient to define a unit vector in the direction of vector $\mathbf{A}$, usually denoted $\hat{a}$, which is defined as

$$\hat{a} = \frac{\mathbf{A}}{|\mathbf{A}|} = (\frac{A_x}{A}, \frac{A_y}{A}, \frac{A_z}{A})$$

In this notation, an alternative way of representing vector $\mathbf{A}$ is $\mathbf{A} = \hat{a}|\mathbf{A}| = \hat{a}A$. Two vectors $\mathbf{A}$ and $\mathbf{B}$ can be added together component by component to produce a third vector:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = \hat{x}(A_x + B_x) + \hat{y}(A_y + B_y) + \hat{z}(A_z + B_z)$$

### 1.1 DOT PRODUCT

The three unit vectors of a given coordinate system are mutually orthogonal, such that the dot product of any two different (orthogonal) unit vectors yields zero (e.g., $\hat{x} \cdot \hat{z} = 0$), while the dot product of a unit vector with itself yields one (e.g., $\hat{z} \cdot \hat{z} = 1$). As a result, the dot product of any two vectors $\mathbf{A}$ and $\mathbf{B}$ yields the scalar quantity

$$\mathbf{A} \cdot \mathbf{B} = A_xB_x + A_yB_y + A_zB_z$$

The dot product is both commutative, such that $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$, and distributive, such that $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$. Note also that $\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}|^2$ and that $A_x = \mathbf{A} \cdot \hat{x}$.

## 1.2 CROSS PRODUCT

The unit vectors of all three coordinate systems are also *right-handed,* such that the cross product of a unit vector with itself yields zero (e.g., $\hat{z} \times \hat{z} = 0$) and the cross product of two orthogonal unit vectors yields a unit vector perpendicular to both following the *right hand rule* (e.g., $\hat{x} \times \hat{y} = \hat{z}$ and $\hat{z} \times \hat{y} = -\hat{x}$). As a result, the cross product between vectors $\mathbf{A}$ and $\mathbf{B}$ is a vector quantity defined (in Cartesian coordinates) in terms of the determinant of the following matrix:

$$\mathbf{A} \times \mathbf{B} = \det \begin{bmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{bmatrix} = \hat{x}(A_y B_z - A_z B_y) + \hat{y}(A_z B_x - A_x B_z) + \hat{z}(A_x B_y - A_y B_x)$$

Note that the cross product operation is not commutative, since $\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}$, nor associative, since $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \neq (\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$, but it is distributive, such that $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$ .
Given three vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, there are two different ways to form a triple product. The scalar triple product is defined as

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})$$

while the vector triple product is defined as $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ and can be expressed more simply as

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$$

## 1.3 CURVILINEAR COORDINATES

The physical properties of electromagnetic fields and waves studied in this class do not depend on the particular coordinate system used to describe the vector quantities. Scalar products, cross products, and other vector operations are all independent of the mathematical frame of reference used to describe them. In practice, however, description and computation of vector quantities in a given problem may be easier and more intuitive in coordinate geometries other than the rectangular Cartesian system. In this class, spherical polar and cylindrical coordinate systems, whose basis unit vector directions depend on location, often will be more useful for describing fields which exhibit spherical and cylindrical symmetries, respectively.



2
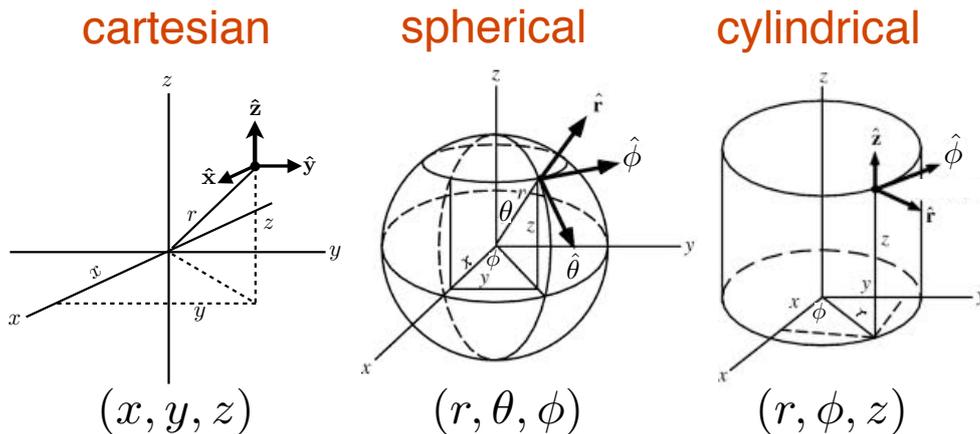
As with Cartesian coordinates, the basis set of three unit vectors in curvilinear coordinate systems are also mutually orthogonal and right-handed. In the spherical coordinate system depicted above, the unit vectors are $(\hat{r}, \hat{\theta}, \hat{\phi})$, where $\hat{r} \times \hat{\theta} = \hat{\phi}$, $r = \sqrt{x^2 + y^2 + z^2}$, and $\theta$ is the polar angle from the $\hat{z}$ axis. The cylindrical coordinate system above uses $(\hat{r}, \hat{\phi}, \hat{z})$, where $r = \sqrt{x^2 + y^2}$, $\phi$ is an azimuthal angle measured in the $\hat{x}\hat{y}$ plane from the $\hat{x}$ axis and the $\hat{z}$ unit vector is the same as that in Cartesian geometry. Transformation between curvilinear and rectangular Cartesian coordinates uses the following trigonometric rules:

| Spherical coordinate variables | Spherical unit vectors |
| --- | --- |
| $r = \sqrt{x^2 + y^2 + z^2}$ | $\hat{r} = \hat{x}\sin\theta\cos\phi + \hat{y}\sin\theta\sin\phi + \hat{z}\cos\theta$ |
| $\theta = \cos^{-1}\left(\dfrac{z}{r}\right)$ | $\hat{\theta} = \hat{x}\cos\theta\cos\phi + \hat{y}\cos\theta\sin\theta - \hat{z}\sin\theta$ |
| $\phi = \tan^{-1}\left(\dfrac{y}{x}\right)$ | $\hat{\phi} = -\hat{x}\sin\phi + \hat{y}\cos\phi$ |

| Cylindrical coordinate variables | Cylindrical unit vectors |
| --- | --- |
| $r = \sqrt{x^2 + y^2}$ | $\hat{r} = \hat{x}\cos\phi + \hat{y}\sin\phi$ |
| $\phi = \tan^{-1}\left(\dfrac{y}{x}\right)$ | $\hat{\phi} = -\hat{x}\sin\phi + \hat{y}\cos\phi$ |
| $z = z$ | $\hat{z} = \hat{z}$ |

## 2. PARTIAL DERIVATIVES

The derivative of a multivariable function, such as the scalar magnitude $A_x(x, y, z, t)$, should be taken one variable at a time while treating the others as constants. For example, the partial derivative of the function $f(x, y) = x^2 y + 4y$ with respect to the variable $x$ is $\dfrac{\partial f}{\partial x} = 2xy$, while $\dfrac{\partial f}{\partial y} = x^2 + 4$.

For higher order partial derivatives with respect to two or more variables, the order of differentiation does not matter. So, for example, the derivative with respect to the variable $y$ of the derivative with respect to $x$ of the function $f(x, y) = x^2 y + 4y$ is:

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}(2xy) = 2x = \frac{\partial}{\partial x}(x^2 + 4) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

The *chain rule* is used to take the derivative of a composition of two or more functions. For example, if $x = \cos(\theta)$ and $\theta = \omega t$, then $\dfrac{\partial x}{\partial t} = \dfrac{\partial x}{\partial \theta}\dfrac{\partial \theta}{\partial t} = -\omega\sin(\omega t)$.

## 2.1 GRADIENT

The vector differential operator, denoted $\nabla$ and called "the del operator", is expressed in Cartesian coordinates as

$$\nabla = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$$

Note that in curvilinear coordinate systems, the del operator takes a different form when acting on the different vector components. When acting on a scalar function $f(x, y, z)$, it yields a vector quantity known as the *gradient* of $f$:

$$\nabla f = \hat{x}\frac{\partial f}{\partial x} + \hat{y}\frac{\partial f}{\partial y} + \hat{z}\frac{\partial f}{\partial z}$$

For example, the linear ramp function $f = 2x$, has a vector gradient $\nabla f = 2\hat{x}$, which has magnitude of 2 along the direction of function variation ($\hat{x}$), independent of position. In contrast, the non-linear scalar function $g = 2xy$ has a position-dependent gradient of $\nabla g = 2y\hat{x} + 2x\hat{y}$. At point $P = (2, 2, 2)$, the gradient of $g$ is $\nabla g|_{(x,y,z)=P} = 4\hat{x} + 4\hat{y}$.

## 2.2 DIVERGENCE

When operating on a vector via the dot product (e.g., $\nabla \cdot \mathbf{A}$) the resulting scalar quantity is known as the *divergence* of vector $\mathbf{A}$:

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

For example, for the vector field $\mathbf{A} = \hat{x}4xy + \hat{y}z^2 + \hat{z}z^3$, the divergence $\nabla \cdot \mathbf{A}$ at the point $P = (2, 2, 2)$ is $\nabla \cdot \mathbf{A}|_{(x,y,z)=P} = 4(2) + 0 + 3(2)^2 = 20$.

## 2.3 CURL

When operating on a vector via the cross product (e.g., $\nabla \times \mathbf{A}$), the resulting vector quantity is known as the *curl* of vector $\mathbf{A}$:

$$\nabla \times \mathbf{A} = \hat{x}(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}) + \hat{y}(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}) + \hat{z}(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y})$$

For example, for the vector field $\mathbf{A} = \hat{x}2y - \hat{y}x$, the curl $\nabla \times \mathbf{A}$ at the point $P = (2, 2, 2)$ is $\nabla \times \mathbf{A}|_{(x,y,z)=P} = \hat{z}(-1 - 2) = -3\hat{z}$.

Fields which are curl-free (fields whose curl evaluates to zero everywhere) are known as *conservative* fields. Any field which can be expressed as the gradient of a scalar function (e.g., $\mathbf{F} = \nabla f$) is by definition a conservative field, since $\nabla \times \nabla f = 0$ for all functions $f$.

## 3. INTEGRATION

The integral of a derivative of a function along a given dimension is simply related to the values of the function at the limits of integration:

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

Similarly, the integral of the gradient of a function (its derivative in three dimensions) along a given three-dimensional path is simply related to the values of the function at the 3D end points of that path, denoted by vectors $\mathbf{A}$ and $\mathbf{B}$:

$$\int_{\mathbf{A}}^{\mathbf{B}} \nabla f \cdot \mathbf{dl} = f(\mathbf{B}) - f(\mathbf{A})$$

Here, $\mathbf{dl} = \hat{x}dx + \hat{y}dy + \hat{z}dz$ is the infinitesimal displacement along the path, and, because the vector field $\nabla f$ is conservative (curl-free) by definition, it is an *exact differential*, such that:

$$\nabla f \cdot \mathbf{dl} = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = df$$

In addition, the value of the line integral is *path-independent*, provided the path end points remain the same. For example, the force $\mathbf{F}$ exerted to move a massive object within a gravitational field is the gradient of potential energy $U$, $\mathbf{F} = \nabla U$, and the work $W$ done to move it from point $\mathbf{A}$ to point $\mathbf{B}$ against gravity, $W = \int_A^B \nabla U \cdot \mathbf{dl} = U(\mathbf{A}) - U(\mathbf{B})$, is independent of the path taken to move the object.

As a result, reversing the limits of integration (integrating backwards along the path) simply yields the negative of the integration along the forward path:

$$\int_{\mathbf{A}}^{\mathbf{B}} \nabla f \cdot \mathbf{dl} = - \int_{\mathbf{B}}^{\mathbf{A}} \nabla f \cdot \mathbf{dl}$$

and thus, for conservative fields, integration around a closed path $C$ (e.g., from point $\mathbf{A}$ to $\mathbf{B}$ and back to $\mathbf{A}$), a quantity known as the *circulation* of $\nabla f$, yields

$$\oint_C \nabla f \cdot \mathbf{dl} = 0$$

### 3.1 CIRCULATION AND STOKES THEOREM

The circulation of an arbitrary vector field $\mathbf{F}$ (that is potentially not conservative and hence potentially has non-zero curl) is related to its curl via the Stokes Theorem:

$$\oint_C \mathbf{F} \cdot \mathbf{dl} = \int_S (\nabla \times \mathbf{F}) \cdot \mathbf{dS}$$

Here, the curl of $\mathbf{F}$ is the circulation of $\mathbf{F}$ *per unit area,* such that the circulation around the closed path $C$ equals the *flux* of $\nabla \times \mathbf{F}$ through any open surface $S$ that is bounded by $C$ (see below). Here, note that $\mathbf{dS}$ is a differential surface area with a vector direction that corresponds to the direction of $\mathbf{dl}$ along $C$ via the right hand rule.

## 3.2 FLUX AND DIVERGENCE THEOREM

The flux of an arbitrary vector field $\mathbf{A}$ through surface $S$, is written as $\int_S \mathbf{A} \cdot \mathbf{dS}$, where $\mathbf{dS} = \hat{n}dS$ is a vector pointing in the direction that is perpendicular to the surface (defined $\hat{n}$ for normal) and $dS$ is a differential area (having magnitude one) on that surface. For example, when integrating in the $+\hat{z}$ direction over a flat surface lying in the $xy$-plane, $\mathbf{dS} = (\hat{x}dx \times \hat{y}dy) = \hat{z}dxdy$. Only the component of the vector field that is parallel to the surface normal direction contributes to flux through the surface, hence the flux calculation utilizes the dot product operation. The flux through a given surface can be calculated from either direction (specified by $\hat{n}$), and a positive flux implies that $\mathbf{A}$ has a component that is parallel to $\hat{n}$ rather than anti-parallel (which yields a negative flux).

When calculating flux through a closed surface $S$, the surface normal vector must point *outwards* from the volume $V$ enclosed by the surface. The flux of a vector field $\mathbf{A}$ through a closed surface is related to the divergence of $\mathbf{A}$ via the divergence theorem:

$$\oint_S \mathbf{A} \cdot \mathbf{dS} = \int_V \nabla \cdot \mathbf{A} \; dV$$

Specifically, the flux of $\mathbf{A}$ through $S$ equals the volume integration of $\nabla \cdot \mathbf{A}$ over the region $V$ enclosed by $S$.