University of Alabama Department of Physics and Astronomy

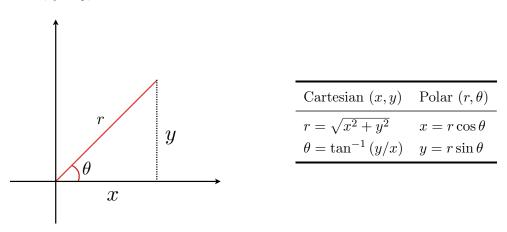
PH 125 / LeClair

Spring 2009

A Short Math Guide

1 Definition of coordinates

Relationship between 2D cartesian (x, y) and polar (r, θ) coordinates. Note that we use the convention that the cartesian unit vectors are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$, rather than $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$, using the substitutions $\hat{\mathbf{x}} = \hat{\mathbf{i}}$, $\hat{\mathbf{y}} = \hat{\mathbf{j}}$, and $\hat{\mathbf{z}} = \hat{\mathbf{k}}$.



Common 3-D coordinate systems and their relationships: cartesian (x, y, z), cylindrical (R, φ, z) , and spherical (r, θ, φ) systems. Note that the cartesian system is a right-handed one – if your fingers point along the x axis $(\hat{\mathbf{x}})$, and your thumb along the y axis $(\hat{\mathbf{y}})$, the z axis $(\hat{\mathbf{z}})$ points out of the back of your hand for a right-handed system.

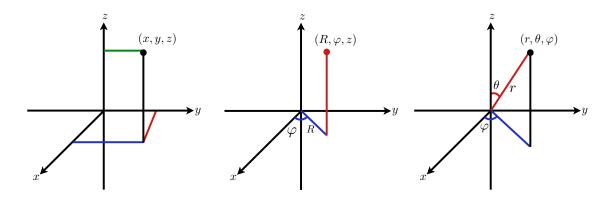


Figure 1: Common 3-D coordinate systems. left: Cartesian (x, y, z), center: cylindrical (R, φ, z) , right: spherical (r, θ, φ) .

Cartesian (x, y, z)	Cylindrical (R,φ,z)	Spherical (r, θ, φ)
$R = \sqrt{x^2 + y^2}$	$x = R\cos\varphi$	$x = r\sin\theta\cos\varphi$
$\varphi = \tan^{-1}\left(\frac{y}{x}\right)$	$y = R\sin\varphi$	$y = r\sin\theta\sin\varphi$
z = z	z = z	$z = r\cos\theta$
$r = \sqrt{x^2 + y^2 + z^2}$	$r = \sqrt{R^2 + z^2}$	$R = r\sin\theta$
$\theta = \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right)$	$\theta = \tan^{-1}\left(\frac{R}{z}\right)$	$\varphi=\varphi$
$\varphi = \tan^{-1}\left(\frac{y}{x}\right)$	$\varphi=\varphi$	$z = r\cos\theta$

Table 1: Relationships between coordinates in different 3D systems.

2 Relationships between unit vectors

Note that these follow directly from the coordinate relationships above plus the normalization condition (i.e., unit vectors must have unit length).

Table 2: Relationships between unit vectors in cartesian and polar systems.

Cartesian (x, y)	Polar (r, θ)
$ \hat{\mathbf{r}} = \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}} \hat{\mathbf{\theta}} = -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}} $	$ \hat{\mathbf{x}} = \sin\theta \hat{\mathbf{r}} + \cos\theta \hat{\mathbf{\theta}} \hat{\mathbf{y}} = \cos\theta \hat{\mathbf{r}} + \sin\theta \hat{\mathbf{\theta}} $

One can also transform between unit vectors using the following matrix equations:

$[\hat{\mathbf{r}}]$	$\cos \theta$	$\sin \theta$	$\begin{bmatrix} \hat{\mathbf{x}} \end{bmatrix}$	$\begin{bmatrix} \hat{\mathbf{x}} \end{bmatrix}$	$\sin \theta$	$\cos \theta \left[\hat{\mathbf{r}} \right]$
$\left[\hat{\theta}\right] =$	$\begin{bmatrix} \cos \theta \\ -\sin \theta \end{bmatrix}$	$\cos \theta$	$\begin{bmatrix} \hat{\mathbf{y}} \end{bmatrix}$	$\begin{bmatrix} \hat{\mathbf{y}} \end{bmatrix}$ =	$\cos \theta$	$ \begin{array}{c} \cos\theta\\ \sin\theta \end{array} \begin{bmatrix} \hat{\mathbf{r}}\\ \hat{\mathbf{\theta}} \end{bmatrix} $

Table 3: Relationships between unit vectors in different 3D systems.

Cartesian (x, y, z)	Cylindrical (R,φ,z)	Spherical (r, θ, φ)
$egin{array}{lll} {\hat {f R}} = rac{x}{R} {\hat {f x}} + rac{y}{R} {\hat {f y}} \ {\hat {f arphi}} \ {\hat {f arphi}} = -rac{y}{R} {\hat {f x}} + rac{x}{R} {\hat {f y}} \end{array}$	· · · · · · · · · · · · · · · · · · ·	$ \hat{\mathbf{x}} = \sin\theta\cos\varphi\hat{\mathbf{r}} + \cos\theta\cos\varphi\hat{\mathbf{\theta}} - \sin\varphi\hat{\boldsymbol{\varphi}} \hat{\mathbf{y}} = \sin\theta\sin\varphi\hat{\mathbf{r}} + \cos\theta\sin\varphi\hat{\mathbf{\theta}} + \cos\varphi\hat{\boldsymbol{\varphi}} $
$egin{array}{lll} m{arphi} & = -rac{1}{R} {f x} + rac{1}{R} {f y} \ {f \hat z} & = {f \hat z} \end{array}$	$\hat{\mathbf{y}} = \sin \phi \mathbf{R} + \cos \phi \phi$ $\hat{\mathbf{z}} = \hat{\mathbf{z}}$	$\mathbf{\hat{z}} = \cos\theta \mathbf{\hat{r}} - \sin\theta \mathbf{\hat{\theta}}$ $\mathbf{\hat{z}} = \cos\theta \mathbf{\hat{r}} - \sin\theta \mathbf{\hat{\theta}}$
$\hat{\mathbf{r}} = \frac{1}{r} \left(x \hat{\mathbf{x}} + y \hat{\mathbf{y}} + z \hat{\mathbf{z}} \right)$	$\mathbf{\hat{r}} = \frac{1}{r} \left(R \mathbf{\hat{R}} + z \mathbf{\hat{z}} \right)$	$\hat{\mathbf{R}} = \sin \theta \hat{\mathbf{r}} + \cos \theta \hat{\mathbf{\theta}}$
$\hat{\boldsymbol{\theta}} = \cos\theta\cos\varphi\hat{\mathbf{x}} + \cos\theta\sin\varphi\hat{\mathbf{y}} - \sin\theta\hat{\mathbf{z}}$	$\hat{\mathbf{\theta}} = \frac{1}{r} \left(z \hat{\mathbf{R}} - R \hat{\mathbf{z}} \right)$	$\hat{oldsymbol{arphi}}=\hat{oldsymbol{arphi}}$
$\hat{oldsymbol{arphi}} = -\sin arphi \hat{f x} + \cos arphi \hat{f y}$	$\hat{oldsymbol{arphi}}=\hat{oldsymbol{arphi}}$	$\mathbf{\hat{z}} = \cos heta \mathbf{\hat{r}} - \sin heta \mathbf{\hat{ heta}}$

The 3D transformations can also be written as matrix equations; for example, transforming from cartesian to spherical polar coordinates, or cartesian and cylindrical coordinates

Î		$\sin\theta\cos\varphi$	$\cos\theta\cos\varphi$	$-\sin \varphi$	$\hat{\mathbf{r}}$	$\begin{bmatrix} \hat{\mathbf{x}} \end{bmatrix}$		$-\sin \varphi$			
ŷ	=	$\sin\theta\sin\varphi$	$\cos\theta\sin\varphi$	$\cos arphi$	θ	$ \hat{\mathbf{y}} =$	$\sin \varphi$	$\cos \varphi$	0	$ \hat{arphi} $	
ĺ		$\cos \theta$	$-\sin \varphi$	0	$\left[\hat{arphi} ight]$	$\begin{bmatrix} \hat{\mathbf{z}} \end{bmatrix}$	0	0	1	$\hat{\mathbf{z}}$	

The remaining transformations can also be written as matrix equations ... which we leave as an exercise to the reader!

3 Rotation of coordinate systems

We will only worry about 2D cartesian (x, y) systems. First, take a normal (x, y) system, and pick a point P(x, y). A clockwise rotation about the origin by θ results in coordinates after rotation P'(x, y):

$$P'_{\theta} = \begin{bmatrix} x'\\y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta\\ +\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x\\y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta\\ +\sin\theta & \cos\theta \end{bmatrix} P$$

Or,

$$x' = x \cos \theta - y \sin \theta$$
$$y' = x \sin \theta + y \cos \theta$$

On the other hand, if make a counterclockwise rotation of θ about the origin (equivalently, a clockwise rotation of $-\theta$), the new coordinates after rotation are

$$P'_{-\theta} = \begin{bmatrix} x'\\y' \end{bmatrix} = \begin{bmatrix} \cos\theta & +\sin\theta\\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x\\y \end{bmatrix} = \begin{bmatrix} \cos\theta & +\sin\theta\\ -\sin\theta & \cos\theta \end{bmatrix} P$$

Or,

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta$$

4 Line, surface, and volume elements in different coordinate systems.

Cartesian

$$d\mathbf{\vec{l}} = \mathbf{\hat{x}} \, dx + \mathbf{\hat{y}} \, dy + \mathbf{\hat{z}} \, dz$$
$$d\mathbf{\vec{S}} = dy \, dz \, \mathbf{\hat{x}} + dx \, dz \, \mathbf{\hat{y}} + dx \, dy \, \mathbf{\hat{z}}$$
$$dV = dx \, dy \, dz$$

Cylindrical

$$d\vec{\mathbf{l}} = \hat{\mathbf{R}} dR + \hat{\boldsymbol{\varphi}} R d\varphi + \hat{\mathbf{z}} dz$$
$$d\vec{\mathbf{S}} = R d\varphi dz \,\hat{\mathbf{R}} + dR dz \,\hat{\boldsymbol{\varphi}} + R dR d\varphi \,\hat{\mathbf{z}}$$
$$dV = R dR d\varphi dz$$

Spherical

$$d\vec{\mathbf{l}} = \hat{\mathbf{r}} dr + \hat{\mathbf{\theta}} r \sin \theta d\theta + \hat{\boldsymbol{\varphi}} r d\varphi$$
$$d\vec{\mathbf{S}} = r^2 \sin \theta d\theta d\varphi \,\hat{\mathbf{r}} + r \sin \theta dr d\varphi \,\hat{\mathbf{\theta}} + r dr d\theta \,\hat{\boldsymbol{\varphi}}$$
$$dV = r^2 \sin \theta dr d\theta d\varphi$$

5 Scalar products

Say you have two vectors, $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$:

$$\vec{\mathbf{a}} = a_x \, \hat{\mathbf{x}} + a_y \, \hat{\mathbf{y}} + a_z \, \hat{\mathbf{z}}$$
$$\vec{\mathbf{b}} = b_x \, \hat{\mathbf{x}} + b_y \, \hat{\mathbf{y}} + b_z \, \hat{\mathbf{z}}$$

The 'dot' or scalar product of these two vectors results in a scalar, or just a number. We find this number by multiplying like components:

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = a_x b_x + a_y b_y + a_z b_z = \sum_{i=1}^n a_i b_i$$

where the latter sum runs over the number of dimensions in your problem (so from 1 to 3 for a normal 3-dimensional problem). The magnitude one of these vectors is simply related to the dot product:

$$|\vec{\mathbf{a}}| = \sqrt{\vec{\mathbf{a}}\cdot\vec{\mathbf{a}}} = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

This gives us a geometric interpretation of the dot prodcut: it is the length of a square of side \vec{a} . We can also write the scalar product in terms of vector magnitudes, and the angle between \vec{a} and \vec{b} :

$$\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = |\vec{\mathbf{a}}| |\vec{\mathbf{b}}| \cos \theta$$

Put another way, given two vectors, the angle between them can be found readily:

$$\theta = \cos^{-1}\left(\frac{\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}}{|\vec{\mathbf{a}}||\vec{\mathbf{b}}|}\right)$$

Of course, this implies that if \vec{a} and \vec{b} are orthogonal (at right angles), their dot product is zero:

if
$$\vec{\mathbf{a}} \perp \vec{\mathbf{b}}$$
, then $\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = 0$

Moreover, two vectors are orthogonal (perpendicular) if and only if their dot product is zero, and they have non-zero length, providing a simple way to test for orthogonality. A few other properties are tabulated below, as well as the scalar product between unit vectors in different coordinate systems.

Table 4: Algebraic properties of the scalar product

formula	relationship
$ec{\mathbf{a}}\cdotec{\mathbf{b}}=ec{\mathbf{b}}\cdotec{\mathbf{a}}$	commutative
$ec{\mathbf{a}} \cdot (ec{\mathbf{b}} + ec{\mathbf{c}}) = ec{\mathbf{a}} \cdot ec{\mathbf{b}} + ec{\mathbf{a}} \cdot ec{\mathbf{c}}$	distributive
$\vec{\mathbf{a}} \cdot (r\vec{\mathbf{b}} + \vec{\mathbf{c}}) = r(\vec{\mathbf{a}} \cdot \vec{\mathbf{b}}) + r(\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})$	bilinear
$(c_1 \vec{\mathbf{a}}) \cdot (c_2 \vec{\mathbf{b}}) = (c_1 c_2) (\vec{\mathbf{a}} \cdot \vec{\mathbf{b}})$	multiplication by scalars
if $\vec{\mathbf{a}} \perp \vec{\mathbf{b}}$, then $\vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = 0$	orthogonality

Table 5: Scalar products of unit vectors

	Ca	rtes	ian	Spherical			Cyl	indrical	
	â	ŷ	â	ŕ	Ô	\hat{arphi}	$\hat{\mathbf{R}}$	\hat{arphi}	ź
â	1	0	0	$\sin\theta\cos\varphi$	$\cos\theta\cos\varphi$	- $\sin \varphi$	$\cos \varphi$	- $\sin \varphi$	0
$\mathbf{\hat{y}}$	0	1	0	$\sin heta \sin arphi$	$\cos heta \sin arphi$	$\cos \varphi$	$\sin arphi$	$\cos \varphi$	0
$\hat{\mathbf{z}}$	0	0	1	$\cos heta$	$-\sin\theta$	0	0	0	1

6 Vector products

The 'cross' or vector product between two polar vectors results in a pseudovector, also known as an 'axial vector.'ⁱ An easy way to remember how to calculate the cross product of these two vectors,

ⁱPseudovectors act just like real vectors, except they gain a sign change under improper rotation. See for example, the Wikipedia page "Pseudovector." An improper rotation is an inversion followed by a normal (proper) rotation, just what we are doing when we switch between right- and left-handed coordinate systems. A proper rotation has no inversion step, just rotation.

 $\vec{\mathbf{c}}\!=\!\vec{\mathbf{a}}\times\vec{\mathbf{b}},$ is to take the determinant of the following matrix:

$$\begin{bmatrix} \mathbf{\hat{x}} & \mathbf{\hat{y}} & \mathbf{\hat{z}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{bmatrix}$$

Or, explicitly:

$$\vec{\mathbf{c}} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \hat{\mathbf{x}} + \begin{vmatrix} a_z & a_x \\ b_z & b_x \end{vmatrix} \hat{\mathbf{y}} + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \hat{\mathbf{z}} = (a_y b_z - a_z b_y) \hat{\mathbf{x}} + (a_z b_x - a_x b_z) \hat{\mathbf{y}} + (a_x b_y - a_y b_x) \hat{\mathbf{z}}$$

One key to remember is that the unit vectors in a orthogonal coordinate system follow a cyclical permutation of xyz. Break the order, and you pick up a minus sign (as that corresponds to an improper rotation, or mirror operation):

 $\begin{aligned} \hat{\mathbf{x}} \times \hat{\mathbf{y}} &= \hat{\mathbf{z}} = -\hat{\mathbf{y}} \times \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \times \hat{\mathbf{z}} &= \hat{\mathbf{x}} = -\hat{\mathbf{z}} \times \hat{\mathbf{y}} \\ \hat{\mathbf{z}} \times \hat{\mathbf{x}} &= \hat{\mathbf{y}} = -\hat{\mathbf{x}} \times \hat{\mathbf{z}} \end{aligned}$

The magnitude of the cross product is

$$|\vec{\mathbf{c}}| = |\vec{\mathbf{a}} \times \vec{\mathbf{b}}| = |\vec{\mathbf{a}}| |\vec{\mathbf{b}}| \sin \theta$$

where θ is the smallest angle between \vec{a} and \vec{b} . Geometrically, the cross product can be viewed as the area of a parallelogram with sides defined by \vec{a} and \vec{b} or the (signed) volume of a parallelepiped defined by the three vectors given. The pseudovector \vec{c} resulting from a cross product of (polar) vectors \vec{a} and \vec{b} is perpendicular to the plane formed by \vec{a} and \vec{b} , with a direction given by the right-hand rule:

$$\vec{\mathbf{a}} \times \vec{\mathbf{b}} = ab\sin\theta\,\hat{\mathbf{n}}$$

where $\hat{\mathbf{n}}$ is a unit vector perpendicular to the plane containing $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$. Note of course that if $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ are collinear (i.e., the angle between them is either 0° or 180°), the cross product is zero. We can use the cross product to define a unit normal vector to a plane formed by vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$:

$$\hat{\mathbf{n}} = \frac{\vec{\mathbf{a}} \times \vec{\mathbf{b}}}{|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|} = \frac{\vec{\mathbf{a}} \times \vec{\mathbf{b}}}{|\vec{\mathbf{a}}||\vec{\mathbf{b}}|\sin\theta} = \frac{\hat{\mathbf{a}} \times \hat{\mathbf{b}}}{\sin\theta}$$

Check this by verifying that the z axis is perpendicular to the x - y plane.

Right-hand rule

- 1. Point the fingers of your right hand along the direction of \vec{a} .
- 2. Point your thumb in the direction of $\vec{\mathbf{b}}$.
- 3. The vector $\vec{\mathbf{c}} = \vec{\mathbf{a}} \times \vec{\mathbf{b}}$ points out from the back of your hand.

The cross product is also anticommutative, distributive over addition, and has numerous other algebraic properties listed in the table below. We also list decompositions of various higher-order products.

Table 6: Algebraic properties of the vector product

formula	relationship
	Telationship
$ec{\mathbf{a}} imes ec{\mathbf{b}} = -ec{\mathbf{b}} imes ec{\mathbf{a}}$	anticommutative
$ec{\mathbf{a}} imes \left(ec{\mathbf{b}} + ec{\mathbf{c}} ight) = \left(ec{\mathbf{a}} imes ec{\mathbf{b}} ight) + \left(ec{\mathbf{a}} imes ec{\mathbf{c}} ight)$	distributive over addition
$(r\vec{\mathbf{a}}) \times \vec{\mathbf{b}} = \vec{\mathbf{a}} \times (r\vec{\mathbf{b}}) = r(\vec{\mathbf{a}} \times \vec{\mathbf{b}})$	compatible with scalar multiplication
$\vec{\mathbf{a}} \times (\vec{\mathbf{b}} \times \vec{\mathbf{c}}) + \vec{\mathbf{b}} \times (\vec{\mathbf{c}} \times \vec{\mathbf{a}}) + \vec{\mathbf{c}} \times (\vec{\mathbf{a}} \times \vec{\mathbf{b}}) = 0$	not associative; obeys Jacobi identity
$ec{\mathbf{a}} imes (ec{\mathbf{b}} imes ec{\mathbf{c}}) = ec{\mathbf{b}} (ec{\mathbf{a}} \cdot ec{\mathbf{b}}) - ec{\mathbf{c}} (ec{\mathbf{a}} \cdot ec{\mathbf{b}})$	triple vector product expansion
$(\vec{\mathbf{a}} imes \vec{\mathbf{b}}) imes \vec{\mathbf{c}} = -\vec{\mathbf{c}} imes (\vec{\mathbf{a}} imes \vec{\mathbf{b}}) = -\vec{\mathbf{a}} (\vec{\mathbf{b}} \cdot \vec{\mathbf{c}}) + \vec{\mathbf{b}} (\vec{\mathbf{a}} \cdot \vec{\mathbf{c}})$	triple vector product expansion
$ec{\mathbf{a}} \cdot (ec{\mathbf{b}} imes ec{\mathbf{c}}) = ec{\mathbf{b}} \cdot (ec{\mathbf{c}} imes ec{\mathbf{a}}) = ec{\mathbf{c}} \cdot (ec{\mathbf{a}} imes ec{\mathbf{b}})$	triple scalar product expansion ^{\dagger}
$ ec{\mathbf{a}} imesec{\mathbf{b}} ^2+ ec{\mathbf{a}}\cdotec{\mathbf{b}} ^2= ec{\mathbf{a}} ^2 ec{\mathbf{b}} ^2$	relation between cross and dot product
$\text{if } \vec{\mathbf{a}} \times \vec{\mathbf{b}} = \vec{\mathbf{a}} \times \vec{\mathbf{c}} \text{ then } \vec{\mathbf{b}} = \vec{\mathbf{c}} \text{ iff } \vec{\mathbf{a}} \cdot \vec{\mathbf{b}} = \vec{\mathbf{a}} \cdot \vec{\mathbf{c}}$	lack of cancellation

[†]Note that the parentheses may be omitted without causing ambiguity, since the dot product cannot be evaluated first. If it were, it would leave the cross product of a vector and a scalar, which is not defined

One higher order product has a particularly nice matrix representation:

$$\vec{\mathbf{a}} \cdot \left(\vec{\mathbf{b}} \times \vec{\mathbf{c}}\right) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} = (a_x b_y c_z - a_x b_z c_y) + (a_y b_z c_x - a_y b_x c_z) + (a_z b_x c_y - a_z b_y c_x) \end{vmatrix}$$

7 Vector Derivatives

7.1 Gradient

In cartesian coordinates, given a scalar function $\varphi(x, y, z)$:

$$\vec{\nabla}\varphi = \frac{\partial\varphi}{\partial x}\hat{x} + \frac{\partial\varphi}{\partial y}\hat{y} + \frac{\partial\varphi}{\partial z}\hat{z}$$
(1)

7.2 Divergence

In cartesian coordinates, given a vector function $\vec{\mathbf{F}}(x, y, z)$:

$$\vec{\nabla} \cdot \vec{\mathbf{F}} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \tag{2}$$

7.3 Curl

In cartesian coordinates, given a vector function $\vec{\mathbf{F}}(x, y, z)$:

$$ec{m{\nabla}} imes ec{m{F}} = egin{bmatrix} \hat{m{x}} & \hat{m{y}} & \hat{m{z}} \ \hat{m{d}} & \hat{m{\partial}} \ \hat{m{\partial}}_{x} & \hat{m{\partial}}_{y} & \hat{m{\partial}}_{z} \ F_{x} & F_{y} & F_{z} \end{bmatrix}$$

8 Example: Solving systems of linear equations

Say we have three equations and three unknowns, and we are left with the pesky problem of solving them. There are many ways to do this, we will illustrate two of them. Take, for example, three equations that result from applying Kirchhoff's rules to a particular multiple loop dc circuit:

 $I_1 - I_2 - I_3 = 0$ $R_1 I_1 + R_3 I_3 = V_1$ $R_2 I_2 - R_3 I_3 = -V_2$

The first way we can proceed is by substituting the first equation into the second:

$$V_1 = R_1 I_1 + R_3 I_3 = R_1 (I_2 + I_3) + R_3 I_3 = R_1 I_2 + (R_1 + R_3) I_3$$

$$\implies V_1 = R_1 I_2 + (R_1 + R_3) I_3$$

Now our three equations look like this:

$$I_1 - I_2 - I_3 = 0$$

$$R_1 I_2 + (R_1 + R_3) I_3 = V_1$$

$$R_2 I_2 - R_3 I_3 = -V_2$$

The last two equations now contain only I_1 and I_2 , so we can solve the third equation for I_2 ...

$$I_2 = \frac{I_3 R_3 - V_2}{R_2}$$

... and plug it in to the second one:

$$V_{1} = R_{1}I_{2} + (R_{1} + R_{3})I_{3} = R_{1}\left(\frac{I_{3}R_{3} - V_{2}}{R_{2}}\right) + (R_{1} + R_{3})I_{3}$$

$$V_{1} + \frac{V_{2}R_{1}}{R_{2}} - \left(R_{1} + R_{3} + \frac{R_{1}R_{3}}{R_{2}}\right)I_{3} = 0$$

$$I_{3} = \frac{V_{1} + \frac{V_{2}R_{1}}{R_{2}}}{R_{1} + R_{3} + \frac{R_{1}R_{3}}{R_{2}}}$$

$$I_{3} = \frac{V_{1}R_{2} + V_{2}R_{1}}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}}$$

Now that you know I_3 , you can plug it in the expression for I_2 above.

What is the second way to solve this? We can start with our original equations, but in a different order:

$$I_1 - I_2 - I_3 = 0$$

$$R_2 I_2 - R_3 I_3 = -V_2$$

$$R_1 I_1 + R_3 I_3 = V_1$$

The trick we want to use is formally known as 'Gaussian elimination,' but it just involves adding these three equations together in different ways to eliminate terms. First, take the first equation above, multiply it by $-R_1$, and add it to the third:

$$\begin{aligned} [-R_1I_1 + R_1I_2 + R_1I_3] &= 0 \\ + & R_1I_1 + R_3I_3 = V_1 \\ \implies & R_1I_2 + (R_1 + R_3)I_3 = V_1 \end{aligned}$$

Now take the second equation, multiply it by $-R_1/R_2$, and add it to the new equation above:

$$-\frac{R_1}{R_2} [R_2 I_2 - R_3 I_3] = -\frac{R_1}{R_2} [-V_2]$$

$$+ \qquad R_1 I_2 + (R_1 + R_3) I_3 = V_1$$

$$\implies \qquad \left(\frac{R_1 R_3}{R_2} + R_1 + R_3\right) I_3 = \frac{R_1}{R_2} V_2 + V_1$$

Now the resulting equation has only I_3 in it. Solve this for I_3 , and proceed as above.

Optional: There is one more way to solve this set of equations using matrices and Cramer's rule,ⁱⁱ if you are familiar with this technique. If you are not familiar with matrices, you can skip to the next problem - you are not required or necessarily expected to know how to do this. First, write the three equations in matrix form:

$$\begin{bmatrix} R_1 & 0 & R_3 \\ 0 & R_2 & -R_3 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} V_1 \\ -V_2 \\ 0 \end{bmatrix}$$
$$\mathbf{aI} = \mathbf{V}$$

The matrix **a** times the column vector **I** gives the column vector **V**, and we can use the determinant of the matrix **a** with Cramer's rule to find the currents. For each current, we construct a new matrix, which is the same as the matrix **a** except that the the corresponding column is replaced the column vector **V**. Thus, for I_1 , we replace column 1 in **a** with **V**, and for I_2 , we replace column 2 in **a** with **V**. We find the current then by taking the new matrix, calculating its determinant, and dividing that by the determinant of **a**. Below, we have highlighted the columns in **a** which have been replaced to make this more clear:

ⁱⁱSee 'Cramer's rule' in the Wikipedia to see how this works.

$$I_{1} = \frac{\begin{vmatrix} V_{1} & 0 & R_{3} \\ -V_{2} & R_{2} & -R_{3} \\ 0 & -1 & -1 \end{vmatrix}}{\det \mathbf{a}} \qquad I_{2} = \frac{\begin{vmatrix} R_{1} & V_{1} & R_{3} \\ 0 & -V_{2} & -R_{3} \\ 1 & 0 & -1 \end{vmatrix}}{\det \mathbf{a}} \qquad I_{3} = \frac{\begin{vmatrix} R_{1} & 0 & V_{1} \\ 0 & R_{2} & -V_{2} \\ 1 & -1 & 0 \end{vmatrix}}{\det \mathbf{a}}$$

Now we need to calculate the determinant of each new matrix, and divide that by the determinant of $\mathbf{a}^{\text{.iii}}$ First, the determinant of \mathbf{a} .

$$\det a = -R_1R_2 - R_1R_3 + 0 - 0 + 0 - R_2R_3 = -(R_1R_2 + R_2R_3 + R_1R_3)$$

We can now find the currents readily from the determinants of the modified matrices above and that of \mathbf{a} we just found:

$$I_{1} = \frac{-V_{1}R_{2} - V_{1}R_{3} + 0 - 0 + V_{2}R_{3} - 0}{-(R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3})} = \frac{V_{1}(R_{2} + R_{3}) - V_{2}R_{3}}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}}$$
$$I_{2} = \frac{R_{1}V_{2} - 0 - V_{1}R_{3} - 0 + 0 + R_{3}V_{2}}{-(R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3})} = \frac{R_{3}V_{1} - V_{2}(R_{1} + R_{3})}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}}$$
$$I_{3} = \frac{0 - R_{1}V_{2} + 0 - 0 + 0 - V_{1}R_{2}}{-(R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3})} = \frac{R_{1}V_{2} + R_{2}V_{1}}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}}$$

These are the same results you would get by continuing on with either of the two previous methods. Both numerically and symbolically, we can see from the above that $I_1 = I_2 + I_3$:

$$I_{2} + I_{3} = \frac{R_{3}V_{1} - V_{2}(R_{1} + R_{3}) + R_{1}V_{2} + R_{2}V_{1}}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}}$$
$$= \frac{V_{1}(R_{2} + R_{2}) + V_{2}(R_{1} - R_{1} - R_{3})}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}}$$
$$= \frac{V_{1}(R_{2} + R_{2}) - V_{2}R_{3}}{R_{1}R_{2} + R_{2}R_{3} + R_{1}R_{3}} = I_{1}$$

ⁱⁱⁱAgain, the Wikipedia entry for 'determinant' is quite instructive.