Vector derivatives

September 7, 2015

In generalizing the idea of a derivative to vectors, we find several new types of object. Here we look at ordinary derivatives, but also the gradient, divergence and curl.

1 Curves

The derivative of a function of a single variable is familiar from calculus, $\frac{df(x)}{dx}$. The simplest generalization of this is when we have a curve in two or three dimensions. A *curve* is a smoothly parameterized sequence of points in space. You are familiar with this idea from mechanics, where a particle's position may be parameterized by time,

$$\mathbf{x}(t) = (x(t), y(t), z(t))$$

At any time t, these three functions give us the position of the moving particle. The complete path of the particle is a "curve".

We can produce a second vector, the velocity vector, from this position vector by differentiation

$$
\begin{aligned}
\mathbf{v}(t) &= \frac{d\mathbf{x}(t)}{dt} \\
&= \left(\frac{dx(t)}{dt}, \frac{dy(t)}{dt}, \frac{dz(t)}{dt} \right)
\end{aligned}
$$

We differentiate each of the three functions with respect to the parameter. This result generalizes to arbitrary curves and parameterizations. If we have a curve parameterized by any parameter λ, $\mathbf{x}(\lambda) = (x(\lambda), y(\lambda), z(\lambda))$, then differentiation gives a tangent vector to the curve at each point,

$$\mathbf{t}(\lambda) = \frac{d\mathbf{x}(\lambda)}{d\lambda}$$

Many parameters can give the same curve. For example,

$$
\begin{aligned}
\mathbf{x}(t) &= (x(t), y(t), z(t)) \\
&= \left(v_{0x}t, y_{0y}t, \frac{1}{2}gt^2 \right)
\end{aligned}
$$

and

$$\mathbf{x}(\sigma) = \left(v_{0x}\sigma^2, y_{0y}\sigma^2, \frac{1}{2}g\sigma^4 \right)$$

describe the same path in space. However, the length of the tangent vector will depend on which parameter is chosen. Thus,

$$\mathbf{v}(t) = \frac{d\mathbf{x}}{dt} = (v_{0x}, v_{0y}, gt)$$

gives the velocity, but (using $\sigma^2 = t$)

$$
\begin{aligned}
\mathbf{t}\left(\sigma\right) &= \frac{d\mathbf{x}}{d\sigma} \\
&= 2\sigma\left(v_{0x}, y_{0y}, g\sigma^2\right) \\
&= 2\sigma\mathbf{v}\left(t\right)
\end{aligned}
$$

is merely *proportional* to the velocity.

2   The gradient

We can produce a vector from a scalar (i.e., a function) by differentiation. Suppose we have a *scalar field*, that is an assignment of a number to each point of Euclidean 3-space:

$$
\phi = \phi\left(x, y, z\right) = \phi\left(\mathbf{x}\right)
$$

This is just the opposite of a curve. Now we have one function of three variables instead of three functions of one variable.

In order to distinguish the different possible derivatives of ϕ, we use a *partial derivative*, holding two of the independent variables constant while we differentiate with respect to the third,

$$
\frac{\partial\phi}{\partial x} = \lim_{\varepsilon\to 0}\frac{\phi\left(x+\varepsilon, y, z\right) - \phi\left(x, y, z\right)}{\varepsilon}
$$

where the expression on the right is the usual definition of the derivative of a function of x.

This means that we have three distinct deravitives of a function, characterizing how it is changing in each of the three coordinate directions. We make a vector of these by combining them with the basis vectors in the corresponding directions. The resulting vector, the *gradient*, gives us the direction in which ϕ is changing the fastest:

$$
\boldsymbol{\nabla}\phi \equiv \hat{\mathbf{i}}\frac{\partial\phi}{\partial x} + \hat{\mathbf{j}}\frac{\partial\phi}{\partial y} + \hat{\mathbf{k}}\frac{\partial\phi}{\partial z} \tag{1}
$$

Notice that the *del operator*, $\boldsymbol{\nabla}$, is written in boldface or with an arrow, $\overrightarrow{\nabla}$ because it results in a vector. Notice that we write $\hat{\mathbf{i}}\frac{\partial\phi}{\partial x} + \hat{\mathbf{j}}\frac{\partial\phi}{\partial y} + \hat{\mathbf{k}}\frac{\partial\phi}{\partial z}$ instead of $\frac{\partial\phi}{\partial x}\hat{\mathbf{i}} + \frac{\partial\phi}{\partial y}\hat{\mathbf{j}} + \frac{\partial\phi}{\partial z}\hat{\mathbf{k}}$ to avoid any confusion about whether the derivative also acts on the basis vectors. This doesn't matter in Cartesian coordinates since the Cartesian basis vectors are constant, but it does matter in other bases. For example, $\frac{\partial}{\partial\theta}\hat{\mathbf{r}}$ does not vanish because the direction of $\hat{\mathbf{r}}$ changes as we vary θ.

We may write del as an operator without necessarily applying it to a function,

$$
\boldsymbol{\nabla} \equiv \hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z} \tag{2}
$$

and, as such, it is a *vector-valued operator*, or simply a vector operator.

Suppose we have a unit vector, $\hat{\mathbf{n}}$, in an arbitrary direction. Using the angular expression for the dot product, the directional derivative of ϕ in the $\hat{\mathbf{n}}$ direction is then

$$
\begin{aligned}
\hat{\mathbf{n}}\cdot\boldsymbol{\nabla}\phi &= |\hat{\mathbf{n}}|\,|\boldsymbol{\nabla}\phi|\cos\theta \\
&= |\boldsymbol{\nabla}\phi|\cos\theta
\end{aligned}
$$

where $|\hat{\mathbf{n}}| = 1$ and where θ is the angle between $\hat{\mathbf{n}}$ and $\boldsymbol{\nabla}\phi$. This expression is clearly largest when $\cos\theta = 1$ and therefore $\theta = 0$. This means that the direction of the vector, $\boldsymbol{\nabla}\phi$, maximizes the directional derivatives of ϕ, thereby giving the direction in which ϕ is changing the fastest.

Example (Problem 1.13a): Find the gradient of the radial coordinate r. We know that $r = \sqrt{x^2 + y^2 + z^2}$ is the magnitude of a radial vector, $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ so the gradient is

$$
\begin{aligned}
\boldsymbol{\nabla} r &= \left(\hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z} \right) \sqrt{x^2 + y^2 + z^2} \\
&= \hat{\mathbf{i}}\frac{\partial \sqrt{x^2 + y^2 + z^2}}{\partial x} + \hat{\mathbf{j}}\frac{\partial \sqrt{x^2 + y^2 + z^2}}{\partial y} + \hat{\mathbf{k}}\frac{\partial \sqrt{x^2 + y^2 + z^2}}{\partial z} \\
&= \frac{1}{2}\frac{1}{\sqrt{x^2 + y^2 + z^2}} \left(\hat{\mathbf{i}}\frac{\partial \left(x^2 + y^2 + z^2\right)}{\partial x} + \hat{\mathbf{j}}\frac{\partial \left(x^2 + y^2 + z^2\right)}{\partial y} + \hat{\mathbf{k}}\frac{\partial \left(x^2 + y^2 + z^2\right)}{\partial z} \right) \\
&= \frac{1}{2}\frac{1}{\sqrt{x^2 + y^2 + z^2}} \left(2x\hat{\mathbf{i}} + 2y\hat{\mathbf{j}} + 2z\hat{\mathbf{k}} \right) \\
&= \frac{1}{r}\mathbf{r} \\
&= \hat{\mathbf{r}}
\end{aligned}
$$

Example: Gravitational potential For example, the gravitational potential energy around a spherical planet falls off as $\frac{1}{r}$,

$$
\Phi = -\frac{GMm}{r}
$$

The negative of the gradient of the potential energy gives the gravitational force,

$$
\begin{aligned}
\mathbf{F} &= -\boldsymbol{\nabla}\Phi \\
&= GMm\boldsymbol{\nabla}\left(\frac{1}{r}\right)
\end{aligned}
$$

We use the chain rule with each derivative,

$$
\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = -\frac{1}{r^2}\frac{\partial r}{\partial x}
$$

so the full gradient is

$$
\begin{aligned}
\mathbf{F} &= -\frac{GMm}{r^2}\boldsymbol{\nabla} r \\
&= -\frac{GMm}{r^2}\hat{\mathbf{r}}
\end{aligned}
$$

3 Divergence

We have seen that we may act with the del operator on a scalar field to produce a vector. We may also apply the del operator to a *vector field* to produce a scalar. A vector field is an assignment of a vector to each point in our space, $\mathbf{v}(x, y, z)$. Define the *divergence* of $\mathbf{v}(x, y, z)$ to be the dot product of the del operator with the vector field,

$$
\boldsymbol{\nabla} \cdot \mathbf{v}(x, y, z) \equiv \left(\hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z} \right) \cdot \left(v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}} \right)
$$

Using the distributive property of the dot product and the product rule of differentiation, we expand the first term

$$
\hat{\mathbf{i}}\frac{\partial}{\partial x} \cdot \left(v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}} \right) = \hat{\mathbf{i}}\frac{\partial}{\partial x} \cdot \left(v_x\hat{\mathbf{i}} \right) + \hat{\mathbf{i}}\frac{\partial}{\partial x} \cdot \left(v_y\hat{\mathbf{j}} \right) + \hat{\mathbf{i}}\frac{\partial}{\partial x} \cdot \left(v_z\hat{\mathbf{k}} \right)
$$

3

$$= \quad \hat{\mathbf{i}} \cdot \frac{\partial}{\partial x}\left(v_x\hat{\mathbf{i}}\right) + \hat{\mathbf{i}} \cdot \frac{\partial}{\partial x}\left(v_y\hat{\mathbf{j}}\right) + \hat{\mathbf{i}} \cdot \frac{\partial}{\partial x}\left(v_z\hat{\mathbf{k}}\right)$$

$$= \quad \left(\hat{\mathbf{i}} \cdot \frac{\partial v_x}{\partial x}\hat{\mathbf{i}} + v_x\hat{\mathbf{i}} \cdot \frac{\partial \hat{\mathbf{i}}}{\partial x}\right) + \left(\hat{\mathbf{i}} \cdot \frac{\partial v_y}{\partial x}\hat{\mathbf{j}} + v_y\hat{\mathbf{i}} \cdot \frac{\partial \hat{\mathbf{j}}}{\partial x}\right) + \left(\frac{\partial v_z}{\partial x}\hat{\mathbf{i}} \cdot \hat{\mathbf{k}} + v_z\hat{\mathbf{i}} \cdot \frac{\partial \hat{\mathbf{k}}}{\partial x}\right)$$

Now, using first the constancy of the Cartesian unit vectors and then the orthogonality of the basis, this reduces to

$$\hat{\mathbf{i}}\frac{\partial}{\partial x} \cdot \left(v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}}\right) \quad = \quad \frac{\partial v_x}{\partial x}\hat{\mathbf{i}} \cdot \hat{\mathbf{i}} + \frac{\partial v_y}{\partial x}\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} + \frac{\partial v_z}{\partial x}\hat{\mathbf{i}} \cdot \hat{\mathbf{k}}$$

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

Similarly the second and third terms of the original expression give us $\frac{\partial v_y}{\partial y}$ and $\frac{\partial v_z}{\partial z}$, respectively. The final result is the sum of the three terms,

$$\boldsymbol{\nabla} \cdot \mathbf{v}\left(x, y, z\right) = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \tag{3}$$

The divergence tells us how much a vector field is increasing as we move outward. The divergence theorem, which we discuss later, quantifies this statement perfectly. In the meantime, consider some examples to get a sense of what is happening.

Divergence of a constant vector field Let a vector field be specified to be the same vector at each point, $\mathbf{v}\left(x, y, z\right) = \mathbf{v}_0 = \left(v_{x0}, v_{y0}, v_{z0}\right)$. Then all of the deriatives vanish, and we have $\boldsymbol{\nabla} \cdot \mathbf{v} = 0$.

Vector field in the x-direction, increasing in the y-direction This also gives zero. We might take the vector field to be

$$\mathbf{v}\left(x, y, z\right) = y^3\hat{\mathbf{i}}$$

The divergence is then

$$\boldsymbol{\nabla} \cdot \mathbf{v}\left(x, y, z\right) \quad = \quad \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$

$$= \quad \frac{\partial}{\partial x}\left(y^3\right)$$

$$= \quad 0$$

Clearly, we need some x-dependence of the x-component, y-dependence of the y, and so on. This means that the vector field grows stronger as we move in the direction the field itself points.

Radial vector One vector that increases in its own direction is the radial vector $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$. Its divergence is

$$\boldsymbol{\nabla} \cdot \mathbf{r} \quad = \quad \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z}$$

$$= \quad 3$$

4   Curl

We may also use the cross product to take derivatives. For this the component form is useful, as long as we keep track of the order of the derivatives and components. Motivatied by our previous result,

$$\mathbf{u} \times \mathbf{v} \quad = \quad \left(u_2v_3 - u_3v_2\right)\hat{\mathbf{i}} + \left(u_3v_1 - u_1v_3\right)\hat{\mathbf{j}} + \left(u_1v_2 - u_2v_1\right)\hat{\mathbf{k}}$$

we replace $(u_x, u_y, u_z) \rightarrow \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$ and define (in Cartesian coordinates):

$$\boldsymbol{\nabla} \times \mathbf{v} \equiv \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \hat{\mathbf{k}} \tag{4}$$

Notice that the curl depends on exactly the derivatives that the divergence does *not* depend on: we have to look at how the components of the vector are changing in the perpendicular directions. Also, we are looking at the *difference* of these derivatives.

To see what is happening, consider a vector field which is everywhere parallel to the xy-plane and independent of the z-direction, so that $v_z = 0$ and all of the z-derivatives vanish,

$$\mathbf{v} = v_x(x,y)\,\hat{\mathbf{i}} + v_y(x,y)\,\hat{\mathbf{j}}$$

The curl is then entirely in the z-direction,

$$\boldsymbol{\nabla} \times \mathbf{v} = \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \hat{\mathbf{k}}$$

The magnitude will be large if v_y increases with increasing x, while v_x decreases with increasing y. This is just what happens if $\mathbf{v}$ is tangent to a circle.

Suppose we have a rotating disk in the xy-plane. If the disk rotates with angular velocity ω, then a point at a distance r from the center at an angle θ from the x-axis has position

$$(x(r,\theta), y(r,\theta)) = (r\cos\theta, r\sin\theta)$$

Uniform rotation means that $\theta = \omega t$, while r remains the same, so the velocity of any point is

$$
\begin{aligned}
\mathbf{v}(r,\theta) &= \frac{d}{dt}\left(r\cos\omega t, r\sin\omega t \right) \\
&= \omega\left(-r\sin\omega t, v\cos\omega t \right) \\
&= \omega\left(-r\sin\theta, r\cos\theta \right) \\
&= \omega\left(-y, x \right)
\end{aligned}
$$

The curl of this vector field is now

$$
\begin{aligned}
\boldsymbol{\nabla} \times \mathbf{v} &\equiv \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \hat{\mathbf{k}} \\
&= -\frac{\partial v_y}{\partial z}\hat{\mathbf{i}} + \frac{\partial v_x}{\partial z}\hat{\mathbf{j}} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \hat{\mathbf{k}} \\
&= -\omega\frac{\partial x}{\partial z}\hat{\mathbf{i}} - \frac{\partial y}{\partial z}\hat{\mathbf{j}} + \left(\frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y} \right) \hat{\mathbf{k}} \\
&= 2\hat{\mathbf{k}}
\end{aligned}
$$

Notice that the direction of the curl is along the axis of the rotation as given by the right hand rule.

5   Combining derivatives

We can combine sums and products of derivatives in various definite ways. First, consider the gradient, which only acts on scalars.

5.1 Gradient

For the gradient acting on a linear combination of functions, we have

$$\boldsymbol{\nabla}\left(af + bg\right) = a\boldsymbol{\nabla}f + b\boldsymbol{\nabla}g$$

for a and b constant. You can easily see this from the usual product rule, since, for example, the $\hat{\mathbf{i}}$ component above is an ordinary partial derivative of a product:

$$
\begin{aligned}
\hat{\mathbf{i}}\frac{\partial}{\partial x}\left(af + bg\right) &= \hat{\mathbf{i}}\left(\frac{\partial}{\partial x}\left(af\right) + \frac{\partial}{\partial x}\left(bg\right)\right) \\
&= \hat{\mathbf{i}}\left(\frac{\partial a}{\partial x}f + a\frac{\partial f}{\partial x} + \frac{\partial b}{\partial x}g + b\frac{\partial g}{\partial x}\right) \\
&= \hat{\mathbf{i}}\left(a\frac{\partial f}{\partial x} + b\frac{\partial g}{\partial x}\right)
\end{aligned}
$$

The generalization to the action of the gradient on a product of functions is immediate,

$$\boldsymbol{\nabla}\left(fg\right) = g\boldsymbol{\nabla}f + f\boldsymbol{\nabla}g$$

We can also ask what happens if the gradient acts on the dot product of two vectors. We can work this out by expanding the dot product in components and using linearity and the product rule,

$$
\begin{aligned}
\boldsymbol{\nabla}\left(\mathbf{u}\cdot\mathbf{v}\right) &= \boldsymbol{\nabla}\left(u_x v_x + u_y v_y + u_z v_z\right) \\
&= \boldsymbol{\nabla}\left(u_x v_x\right) + \boldsymbol{\nabla}\left(u_y v_y\right) + \boldsymbol{\nabla}\left(u_z v_z\right) \\
&= \left(\boldsymbol{\nabla}u_x\right)v_x + u_x\boldsymbol{\nabla}v_x + \left(\boldsymbol{\nabla}u_y\right)v_y + u_y\boldsymbol{\nabla}v_y + \left(\boldsymbol{\nabla}u_z\right)v_z + u_z\boldsymbol{\nabla}v_z
\end{aligned}
$$

This is correct, but cumbersome. We do better if we use some index notation to write either

$$
\begin{aligned}
\boldsymbol{\nabla}\left(\mathbf{u}\cdot\mathbf{v}\right) &= \boldsymbol{\nabla}\left(\sum_{i=1}^{3}u_i v_i\right) \\
&= \sum_{i=1}^{3}\left(\boldsymbol{\nabla}u_i\right)v_i + \sum_{i=1}^{3}u_i\boldsymbol{\nabla}v_i
\end{aligned}
$$

or, perhaps more clearly

$$
\begin{aligned}
\nabla_k\left(\mathbf{u}\cdot\mathbf{v}\right) &= \nabla_k\left(\sum_{i=1}^{3}u_i v_i\right) \\
&= \left(\nabla_k\mathbf{u}\right)\cdot\mathbf{v} + \mathbf{u}\cdot\nabla_k\mathbf{v}
\end{aligned}
$$

Any of these three makes it clear what the final vector components are. We might also reconstruct the vector, writing,

$$\boldsymbol{\nabla}\left(\mathbf{u}\cdot\mathbf{v}\right) = \hat{\mathbf{i}}\left[\frac{\partial\mathbf{u}}{\partial x}\cdot\mathbf{v} + \mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial x}\right] + \hat{\mathbf{j}}\left[\frac{\partial\mathbf{u}}{\partial y}\cdot\mathbf{v} + \mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial y}\right] + \hat{\mathbf{k}}\left[\frac{\partial\mathbf{u}}{\partial z}\cdot\mathbf{v} + \mathbf{u}\cdot\frac{\partial\mathbf{v}}{\partial z}\right]$$

We may also take the gradient of a divergence,

$$\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\mathbf{v}\right)$$

Since the divergence results in a function, its gradient is a vector.

5.2   Divergence

The divergence is also linear. For a constant linear combination of vectors,

$$a\mathbf{u} + b\mathbf{v}$$

the divergence distributes normally,

$$\boldsymbol{\nabla} \cdot (a\mathbf{u} + b\mathbf{v}) \quad = \quad a\boldsymbol{\nabla} \cdot \mathbf{u} + b\boldsymbol{\nabla} \cdot \mathbf{v}$$

If we multiply a vector field by a function, $f\mathbf{v} = (fv_x, fv_y, fv_z)$, we get derivatives of that function as well:

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot (f\mathbf{v}) \quad &= \quad \frac{\partial (fv_x)}{\partial x} + \frac{\partial (fv_y)}{\partial y} + \frac{\partial (fv_z)}{\partial z} \\
&= \quad \frac{\partial f}{\partial x}v_x + \frac{\partial f}{\partial y}v_y + \frac{\partial f}{\partial z}v_z + f\frac{\partial v_x}{\partial x} + f\frac{\partial v_y}{\partial y} + f\frac{\partial v_z}{\partial z} \\
&= \quad (\boldsymbol{\nabla} f) \cdot \mathbf{v} + f\boldsymbol{\nabla} \cdot \mathbf{v}
\end{aligned}
$$

Notice that the answer can still be written in terms of the gradient and divergence.

Since the cross product of two vectors is a vector, we may consider its divergence. Working through the components, we have

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot (\mathbf{u} \times \mathbf{v}) \quad &= \quad \boldsymbol{\nabla} \cdot \left[(u_2 v_3 - u_3 v_2)\,\hat{\mathbf{i}} + (u_3 v_1 - u_1 v_3)\,\hat{\mathbf{j}} + (u_1 v_2 - u_2 v_1)\,\hat{\mathbf{k}} \right] \\
&= \quad \frac{\partial}{\partial x}(u_2 v_3 - u_3 v_2) + \frac{\partial}{\partial y}(u_3 v_1 - u_1 v_3) + \frac{\partial}{\partial z}(u_1 v_2 - u_2 v_1) \\
&= \quad \left(\frac{\partial u_2}{\partial x}v_3 - \frac{\partial u_3}{\partial x}v_2 \right) + \left(\frac{\partial u_3}{\partial y}v_1 - \frac{\partial u_1}{\partial y}v_3 \right) + \left(\frac{\partial u_1}{\partial z}v_2 - \frac{\partial u_2}{\partial z}v_1 \right) \\
&\quad + \left(u_2 \frac{\partial v_3}{\partial x} - u_3 \frac{\partial v_2}{\partial x} \right) + \left(u_3 \frac{\partial v_1}{\partial y} - u_1 \frac{\partial v_3}{\partial y} \right) + \left(u_1 \frac{\partial v_2}{\partial z} - u_2 \frac{\partial v_1}{\partial z} \right) \\
&= \quad v_1 \left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z} \right) + v_2 \left(\frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x} \right) + v_3 \left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \right) \\
&\quad - u_1 \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z} \right) - u_2 \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x} \right) - u_3 \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) \\
&= \quad \mathbf{v} \cdot (\boldsymbol{\nabla} \times \mathbf{u}) - \mathbf{u} \cdot (\boldsymbol{\nabla} \times \mathbf{v})
\end{aligned}
$$

We can also take the divergence of the curl, and the result is simple and important,

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) \quad &= \quad \boldsymbol{\nabla} \cdot \left(\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right)\hat{\mathbf{i}} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right)\hat{\mathbf{j}} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right)\hat{\mathbf{k}} \right) \\
&= \quad \frac{\partial}{\partial x}\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) + \frac{\partial}{\partial y}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) + \frac{\partial}{\partial z}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \\
&= \quad \frac{\partial^2 v_z}{\partial x \partial y} - \frac{\partial^2 v_y}{\partial x \partial z} + \frac{\partial^2 v_x}{\partial y \partial z} - \frac{\partial^2 v_z}{\partial y \partial x} + \frac{\partial^2 v_y}{\partial z \partial x} - \frac{\partial^2 v_x}{\partial z \partial y} \\
&= \quad \left(\frac{\partial^2 v_z}{\partial x \partial y} - \frac{\partial^2 v_z}{\partial y \partial x} \right) + \left(\frac{\partial^2 v_x}{\partial y \partial z} - \frac{\partial^2 v_x}{\partial z \partial y} \right) + \left(\frac{\partial^2 v_y}{\partial z \partial x} - \frac{\partial^2 v_y}{\partial x \partial z} \right)
\end{aligned}
$$

and because mixed partial derivatives always commute, e.g., $\frac{\partial^2 v_z}{\partial x \partial y} = \frac{\partial^2 v_z}{\partial y \partial x}$ this vanishes identically. *The divergence of the curl of any vector field is zero:*

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) = 0 \tag{5}$$

This makes sense because the curl isolates the amount that the vector field is circling around, which is orthogonal to how it grows along itself.

Finally, we derive an important new operator, the *Laplacian,* if we combine the divergence with the gradient,

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

$$= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

This operator is important enough to have its own symbol:

$$\nabla^2 f \equiv \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} f) \tag{6}$$

The Laplacian operator, ∇^2 is a scalar, so it need not be bold.

5.3   Curl

Finally, we consider the curl of every vector in sight. For a constant linear combination of vectors, the answer is just

$$\boldsymbol{\nabla} \times (a\mathbf{u} + b\mathbf{v}) = a\boldsymbol{\nabla} \times \mathbf{u} + b\boldsymbol{\nabla} \times \mathbf{v}$$

but once again, if we have a function times a vector, we get derivatives of the functions,

$$\boldsymbol{\nabla} \times (f\mathbf{v}) = \left(\frac{\partial (fv_z)}{\partial y} - \frac{\partial (fv_y)}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial (fv_x)}{\partial z} - \frac{\partial (fv_z)}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial (fv_y)}{\partial x} - \frac{\partial (fv_x)}{\partial y} \right) \hat{\mathbf{k}}$$

Separating the derivatives of f from the curl of $\mathbf{v}$ as we did with the divergence, we find a cross product in addition to f times the curl of $\mathbf{v}$

$$\boldsymbol{\nabla} \times (f\mathbf{v}) = f\boldsymbol{\nabla} \times \mathbf{v} + (\boldsymbol{\nabla} f) \times \mathbf{v} \tag{7}$$

The curl of a gradient becomes,

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

$$= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y} \right) \hat{\mathbf{i}} + \left(\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z} \right) \hat{\mathbf{j}} + \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) \hat{\mathbf{k}}$$

Everything cancels because of the equality of mixed partials and we have

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) \equiv 0 \tag{8}$$

for any function f.

There are two further identities to mention. The curl of a cross product is a bit messy,

$$\boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{w} \cdot \boldsymbol{\nabla}) \mathbf{v} + (\boldsymbol{\nabla} \cdot \mathbf{w}) \mathbf{v} - (\boldsymbol{\nabla} \cdot \mathbf{v}) \mathbf{w} - (\mathbf{v} \cdot \boldsymbol{\nabla}) \mathbf{w}$$

but the curl of a curl is simple and useful to remember:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v}) = \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \mathbf{v}) - \nabla^2 \mathbf{v} \tag{9}$$

These can be proved by writing out the components as for the previous results, but the answers follow more quickly if we introduce a new object, the Levi-Civita tensor. Griffiths doesn't introduce this until problem 6.22, but I'll digress to describe it here. You won't be required to use it.

6 The Levi-Civita tensor (advanced)

This section is more advanced. It is not required, but the techniques are extremely useful for proving identities involving cross products and curls.

In 3-dimensions, we define the Levi-Civita tensor, ε_{ijk}, to be totally antisymmetric, so we get a minus sign under interchange of any pair of indices. We work throughout in Cartesian coordinate. This means that most of the 27 components are zero, since, for example,

$$\varepsilon_{212} = -\varepsilon_{212} = 0$$

if we imagine interchanging the two 2s. This means that the only nonzero components are the ones for which i, j and k all take different value. There are only six of these, and all of their values are determined once we choose any one of them. Define

$$\varepsilon_{123} \equiv 1$$

Then by antisymmetry it follows that

$$
\begin{aligned}
\varepsilon_{123} &= \varepsilon_{231} = \varepsilon_{312} = +1 \\
\varepsilon_{132} &= \varepsilon_{213} = \varepsilon_{321} = -1
\end{aligned}
$$

All other components are zero.

Using ε_{ijk} we can write index expressions for the cross product and curl. The ith component of the cross product is given by

$$[\mathbf{u} \times \mathbf{v}]_i = \varepsilon_{ijk} u_j v_k$$

as we check by simply writing out the sums for each value of i,

$$
\begin{aligned}
[\mathbf{u} \times \mathbf{v}]_1 &= \varepsilon_{1jk} u_j v_k \\
&= \varepsilon_{123} u_2 v_3 + \varepsilon_{132} u_3 v_2 + (all\ other\ terms\ are\ zero) \\
&= u_2 v_3 - u_3 v_2 \\
[\mathbf{u} \times \mathbf{v}]_2 &= \varepsilon_{2jk} u_j v_k \\
&= \varepsilon_{231} u_3 v_1 + \varepsilon_{213} u_1 v_3 \\
&= u_3 v_1 - u_1 v_3 \\
[\mathbf{u} \times \mathbf{v}]_3 &= \varepsilon_{3jk} u_j v_k \\
&= u_1 v_2 - u_2 v_1
\end{aligned}
$$

We get the curl simply by replacing u_i by $\nabla_i = \frac{\partial}{\partial x_i}$,

$$[\boldsymbol{\nabla} \times \mathbf{v}]_i = \varepsilon_{ijk} \nabla_j v_k$$

If we sum these expressions with basis vectors $\mathbf{e}_i$, where $\mathbf{e}_1 = \mathbf{i}, \mathbf{e}_2 = \mathbf{j}, \mathbf{e}_3 = \mathbf{k}$, we may write these as vectors:

$$
\begin{aligned}
\mathbf{u} \times \mathbf{v} &= [\mathbf{u} \times \mathbf{v}]_i \, \mathbf{e}_i \\
&= \varepsilon_{ijk} u_j v_k \mathbf{e}_i \\
\boldsymbol{\nabla} \times \mathbf{v} &= \varepsilon_{ijk} \mathbf{e}_i \nabla_j v_k
\end{aligned}
$$

There are useful identities involving pairs of Levi-Civita tensors. The most general is

$$\varepsilon_{ijk} \varepsilon_{lmn} = \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{kn}$$

To check this, first notice that the right side is antisymmetric in i, j, k and antisymmetric in l, m, n. For example, if we interchange i and j, we get

$$\varepsilon_{jik} \varepsilon_{lmn} = \delta_{jl}\delta_{im}\delta_{kn} + \delta_{jm}\delta_{in}\delta_{kl} + \delta_{jn}\delta_{il}\delta_{km} - \delta_{jl}\delta_{in}\delta_{km} - \delta_{jn}\delta_{im}\delta_{kl} - \delta_{jm}\delta_{il}\delta_{kn}$$

9

Now interchange the first pair of Kronecker deltas in each term, to get i, j, k in the original order, then rearrange terms, then pull out an overall sign,

$$
\begin{aligned}
\varepsilon_{jik}\varepsilon_{lmn} &= \delta_{im}\delta_{jl}\delta_{kn} + \delta_{in}\delta_{jm}\delta_{kl} + \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jl}\delta_{km} - \delta_{im}\delta_{jn}\delta_{kl} - \delta_{il}\delta_{jm}\delta_{kn} \\
&= -\delta_{il}\delta_{jm}\delta_{kn} - \delta_{im}\delta_{jn}\delta_{kl} - \delta_{in}\delta_{jl}\delta_{km} + \delta_{il}\delta_{jn}\delta_{km} + \delta_{in}\delta_{jm}\delta_{kl} + \delta_{im}\delta_{jl}\delta_{kn} \\
&= -\left(\delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{kn}\right) \\
&= -\varepsilon_{ijk}\varepsilon_{lmn}
\end{aligned}
$$

Total antisymmetry means that if we know one component, the others are all determined uniquely. Therefore, set $i = l = 1, j = m = 2, k = n = 3$, to see that

$$
\begin{aligned}
\varepsilon_{123}\varepsilon_{123} &= \delta_{11}\delta_{22}\delta_{33} + \delta_{12}\delta_{23}\delta_{31} + \delta_{13}\delta_{21}\delta_{32} - \delta_{11}\delta_{23}\delta_{32} - \delta_{13}\delta_{22}\delta_{31} - \delta_{12}\delta_{21}\delta_{33} \\
&= \delta_{11}\delta_{22}\delta_{33} \\
&= 1
\end{aligned}
$$

Check one more case. Let $i = 1, j = 2, k = 3$ again, but take $l = 3, m = 2, n = 1$. Then we have

$$
\begin{aligned}
\varepsilon_{123}\varepsilon_{321} &= \delta_{13}\delta_{22}\delta_{31} + \delta_{12}\delta_{21}\delta_{33} + \delta_{11}\delta_{23}\delta_{32} - \delta_{13}\delta_{21}\delta_{32} - \delta_{11}\delta_{22}\delta_{33} - \delta_{12}\delta_{23}\delta_{31} \\
&= -\delta_{11}\delta_{22}\delta_{33} \\
&= -1
\end{aligned}
$$

as expected.

We get a second identity by setting $n = k$ and summing,

$$
\begin{aligned}
\varepsilon_{ijk}\varepsilon_{lmk} &= \delta_{il}\delta_{jm}\delta_{kk} + \delta_{im}\delta_{jk}\delta_{kl} + \delta_{ik}\delta_{jl}\delta_{km} - \delta_{il}\delta_{jk}\delta_{km} - \delta_{ik}\delta_{jm}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{kk} \\
&= 3\delta_{il}\delta_{jm} + \delta_{im}\delta_{jl} + \delta_{im}\delta_{jl} - \delta_{il}\delta_{jm} - \delta_{il}\delta_{jm} - 3\delta_{im}\delta_{jl} \\
&= (3 - 1 - 1)\delta_{il}\delta_{jm} - (3 - 1 - 1)\delta_{im}\delta_{jl} \\
&= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}
\end{aligned}
$$

so we have a much simpler, and very useful, relation

$$
\varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}
$$

A second sum gives another identity. Setting $m = j$ and summing again,

$$
\begin{aligned}
\varepsilon_{ijk}\varepsilon_{ljk} &= \delta_{il}\delta_{mm} - \delta_{im}\delta_{ml} \\
&= 3\delta_{il} - \delta_{il} \\
&= 2\delta_{il}
\end{aligned}
$$

Setting the last two indices equal and summing provides a check on our normalization,

$$
\varepsilon_{ijk}\varepsilon_{ijk} = 2\delta_{ii} = 6
$$

This is correct, since there are only six nonzero components and we are summing their squares.

Collecting these results,

$$
\begin{aligned}
\varepsilon_{ijk}\varepsilon_{lmn} &= \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{im}\delta_{jl}\delta_{kn} \\
\varepsilon_{ijk}\varepsilon_{lmk} &= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \\
\varepsilon_{ijk}\varepsilon_{ljk} &= 2\delta_{il} \\
\varepsilon_{ijk}\varepsilon_{ijk} &= 6
\end{aligned}
$$

Now we use these properties to prove some vector identities. First, consider the triple product,

$$
\begin{aligned}
\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= u_i \left[\mathbf{v} \times \mathbf{w} \right]_i \\
&= u_i \varepsilon_{ijk} v_j w_k \\
&= \varepsilon_{ijk} u_i v_j w_k
\end{aligned}
$$

Because $\varepsilon_{ijk} = \varepsilon_{kij} = \varepsilon_{jki}$, we may write this in two other ways,

$$
\begin{aligned}
\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \varepsilon_{ijk} u_i v_j w_k \\
&= \varepsilon_{kij} u_i v_j w_k \\
&= w_k \varepsilon_{kij} u_i v_j \\
&= w_i \left[\mathbf{u} \times \mathbf{v} \right]_i \\
&= \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \varepsilon_{ijk} u_i v_j w_k \\
&= \varepsilon_{jki} u_i v_j w_k \\
&= v_j \left[\mathbf{w} \times \mathbf{u} \right]_j \\
&= \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u})
\end{aligned}
$$

so that we have established

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

and we get the negative permutations by interchanging the order of the vectors in the cross products.

Next, consider a double cross product:

$$
\begin{aligned}
\left[\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) \right]_i &= \varepsilon_{ijk} u_j \left[\mathbf{v} \times \mathbf{w} \right]_k \\
&= \varepsilon_{ijk} u_j \varepsilon_{klm} v_l w_m \\
&= \varepsilon_{ijk} \varepsilon_{klm} u_j v_l w_m \\
&= \varepsilon_{ijk} \varepsilon_{lmk} u_j v_l w_m \\
&= \left(\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \right) u_j v_l w_m \\
&= \delta_{il} \delta_{jm} u_j v_l w_m - \delta_{im} \delta_{jl} u_j v_l w_m \\
&= \left(\delta_{il} v_l \right) \left(\delta_{jm} u_j w_m \right) - \left(\delta_{jl} u_j v_l \right) \left(\delta_{im} w_m \right) \\
&= v_i \left(u_m w_m \right) - \left(u_j v_j \right) w_i
\end{aligned}
$$

Returning to vector notation, this is the $BAC\text{-}CAB$ rule,

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

Finally, look at the curl of a cross product,

$$
\begin{aligned}
\left[\boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{w}) \right]_i &= \varepsilon_{ijk} \nabla_j \left[\mathbf{v} \times \mathbf{w} \right]_k \\
&= \varepsilon_{ijk} \nabla_j \left(\varepsilon_{klm} v_l w_m \right) \\
&= \varepsilon_{ijk} \varepsilon_{klm} \left((\nabla_j v_l) w_m + v_l \nabla_j w_m \right) \\
&= \left(\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \right) \left((\nabla_j v_l) w_m + v_l \nabla_j w_m \right) \\
&= \delta_{il} \delta_{jm} \left((\nabla_j v_l) w_m + v_l \nabla_j w_m \right) - \delta_{im} \delta_{jl} \left((\nabla_j v_l) w_m + v_l \nabla_j w_m \right) \\
&= (\nabla_m v_i) w_m + v_i \nabla_m w_m - (\nabla_j v_j) w_i - v_j \nabla_j w_i
\end{aligned}
$$

Restoring the vector notation, we have

$$
\boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{w} \cdot \boldsymbol{\nabla}) \mathbf{v} + (\boldsymbol{\nabla} \cdot \mathbf{w}) \mathbf{v} - (\boldsymbol{\nabla} \cdot \mathbf{v}) \mathbf{w} - (\mathbf{v} \cdot \boldsymbol{\nabla}) \mathbf{w}
$$

If you doubt the advantages here, try to prove these identities by explicitly writing out all of the components!

7 Exercises (required)

Work the following problems from Griffiths, Chapter 1 (page numbers refer to the 3rd edition):

- Problem 15 (page 18)

- Problem 18 (page 20)

- Problem 25 (page 24)