Gradient, Divergence, Laplacian, and Curl in Non-Euclidean Coordinate Systems

Math 225 supplement to Colley's text, Section 3.4

Many problems are more easily stated and solved using a coordinate system other than rectangular coordinates, for example polar coordinates. It is convenient to have formulas for gradients and Laplacians of functions and divergence and curls of vector fields in terms of other coordinate systems.

Coordinate Vector Fields

To do this we need analogs of $\mathbf{i}$ and $\mathbf{j}$ that are adapted to other coordinate systems. These are called *coordinate vector fields.*

What are the important properties of $\mathbf{i}$ and $\mathbf{j}$ relative to the (x, y) coordinate system?

A) They are orthogonal.

B) They are unit vectors.

C) $\mathbf{i}$ points in the direction in which x increases and y is constant, and $\mathbf{j}$ points in the direction in which y increases and x is constant.

D) $\mathbf{i}$ and $\mathbf{j}$ represent differentiation of a function by x and y. What this means is that if f is a differentiable function, then

$$\frac{\partial f}{\partial x} = D_{\mathbf{i}} f \qquad \text{and} \qquad \frac{\partial f}{\partial y} = D_{\mathbf{j}} f.$$

The function f plays a minor role in these equations. It can be replaced by any differentiable quantity that depends on x and y. To deemphasize its role, we can write the equations as

$$\frac{\partial}{\partial x} = D_{\mathbf{i}} \qquad \text{and} \qquad \frac{\partial}{\partial y} = D_{\mathbf{j}},$$

which are equations about differential operators.

To indicate that $\mathbf{i}$ and $\mathbf{j}$ are related to the (x, y) coordinate system, we could write them as $\mathbf{v}_x$ and $\mathbf{v}_y$ instead, or as $\mathbf{u}_x$ and $\mathbf{u}_y$ to remind ourselves that they are unit vectors. As a hint of things to come, if we think of $X = (x, y)$ as a moving point depending on x and y, we have

$$\mathbf{i} = \mathbf{v}_x = \frac{\partial X}{\partial x} \qquad \text{and} \qquad \mathbf{j} = \mathbf{v}_y = \frac{\partial X}{\partial y}. \tag{1}$$

Coordinate Vector Fields for Polar Coordinates. Now suppose we have a polar coordinate system (r, θ) in the plane E^2. What we need are vectors $\mathbf{v}_r$ and $\mathbf{v}_\theta$ that are related to polar coordinates in the same way that $\mathbf{i}$ and $\mathbf{j}$ are related to Cartesian coordinates. These will be the coordinate vector fields for polar coordinates. Unfortunately, for an arbitrary coordinate system it is *impossible* to satisfy all four of the properties above. (It is a fairly deep theorem in differential geometry that the *only* coordinate systems with coordinate vector fields that satisfy all four properties are Euclidean coordinates.) It turns out that the important ones are C) and D), since these specify how the vectors are related to computing the partial derivatives of a function.

When we use polar coordinates, the position X is a function of r and θ, that is, $X = X(r, \theta)$. Taking our cue from (1), we define

$$\mathbf{v}_r = \frac{\partial X}{\partial r} \qquad \text{and} \qquad \mathbf{v}_\theta = \frac{\partial X}{\partial \theta}. \tag{2}$$

We can verify that these satisfy C) and D) for polar coordinates without even needing their formulas in terms of $\mathbf{i}$ and $\mathbf{j}$.

To see that they satisfy C), note that the curve $\theta \mapsto X(r, \theta)$ is, by definition, the path taken when θ varies and r is fixed, namely a circle centered at the origin, oriented counterclockwise. Thus, its velocity $\mathbf{v}_\theta = \frac{\partial X}{\partial \theta}$ points in the direction in which θ increases and r is fixed. Similarly, $\mathbf{v}_r = \frac{\partial X}{\partial r}$ points in the direction in which r increases and θ is fixed.

To verify D) we need the function-curve version of the chain rule, which we saw earlier in the course: if f is differentiable function on E^n and $X(t)$ is the position of a differentiable curve in E^n, then

$$\frac{d}{dt} f\big(X(t)\big) = \nabla f_{X(t)} \cdot X'(t).$$

If it is implicitly understood that X is a function of t, we might write this as $\frac{d}{dt} f(X) = \nabla f_X \cdot \frac{dX}{dt}$, or even as $\frac{df}{dt} = \nabla f \cdot \frac{dX}{dt}$, although the latter form is an abuse of notation since f is not really a function of t.

Now suppose $f : E^2 \to \mathbb{R}$ is differentiable. Then f is a function of r and θ (a formula for f may or may not be written in terms of r and θ, but we can still talk about f as a function of r and θ since they determine position). If we hold θ fixed, let r vary, and use the chain rule formula, we get

$$\frac{\partial f}{\partial r} = \frac{\partial}{\partial r} f\big(X(r, \theta)\big) = \nabla f_{X(r,\theta)} \cdot \frac{\partial X}{\partial r}(r, \theta) = \nabla f \cdot \mathbf{v}_r = D_{\mathbf{v}_r} f.$$

Similarly, if we hold r fixed and let θ vary, we get $\frac{\partial f}{\partial \theta} = \nabla f \cdot \frac{\partial X}{\partial \theta} = \nabla f \cdot \mathbf{v}_\theta = D_{\mathbf{v}_\theta} f$. The equations

$$\frac{\partial f}{\partial r} = D_{\mathbf{v}_r} f \qquad \text{and} \qquad \frac{\partial f}{\partial \theta} = D_{\mathbf{v}_\theta} f$$

say that $\mathbf{v}_r$ and $\mathbf{v}_\theta$ represent differentiation of a function by r and θ, which verifies D). To deemphasize the role of f, we can write these as equations of differential operators:

$$\frac{\partial}{\partial r} = D_{\mathbf{v}_r} \qquad \text{and} \qquad \frac{\partial}{\partial \theta} = D_{\mathbf{v}_\theta}.$$

To check A) and B) we need formulas for $\mathbf{v}_r$ and $\mathbf{v}_\theta$ in terms of the related Euclidean coordinate system given by $(x, y) = (r\cos\theta, r\sin\theta)$. We have

$$\mathbf{v}_r = \frac{\partial X}{\partial r} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j} \qquad \text{and} \qquad \mathbf{v}_\theta = \frac{\partial X}{\partial \theta} = -r\sin\theta\,\mathbf{i} + r\cos\theta\,\mathbf{j}.$$

From these it is easily verified that $\mathbf{v}_r$ and $\mathbf{v}_\theta$ are orthogonal, so they satisfy A). However, only one of them is a unit vector, namely $\mathbf{v}_r$, so they do not satisfy B). We see that $\mathbf{v}_r$ is a unit vector pointing directly away from the origin. We have $\|\mathbf{v}_\theta\| = r$. More specifically, we have that $\mathbf{v}_\theta = r(\mathbf{v}_r)_\perp$.

Since unit vectors are useful, we define

$$\mathbf{u}_r = \mathbf{v}_r = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j} \qquad \text{and} \qquad \mathbf{u}_\theta = \frac{\mathbf{v}_\theta}{\|\mathbf{v}_\theta\|} = \frac{\mathbf{v}_\theta}{r} = -\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j},$$

and note that $\mathbf{u}_\theta = (\mathbf{u}_r)_\perp$. Colley neglected to give these formulas in the section on polar coordinates (pg. 62), but they are the same as two of the three coordinate vector fields for cylindrical coordinates on page 71. You should verify the coordinate vector field formulas for spherical coordinates on page 72.

For any differentiable function f we have

$$D_{\mathbf{u}_r} f = D_{\mathbf{v}_r} f = \frac{\partial f}{\partial r} \qquad \text{and} \qquad D_{\mathbf{u}_\theta} f = \frac{1}{r} D_{\mathbf{v}_\theta} f = \frac{1}{r}\frac{\partial f}{\partial \theta}. \tag{3}$$

The second formula follows by applying the property $D_{a\mathbf{v}} f = \nabla f \cdot (a\mathbf{v}) = a\nabla f \cdot \mathbf{v} = aD_{\mathbf{v}} f$ to the case when $a\mathbf{v} = \frac{1}{r}\mathbf{v}_\theta$. This illustrates an important feature of non-Euclidean coordinate systems that is often counterintuitive:

> *The directional derivative of a function in the direction of a unit coordinate vector need not be equal to the partial derivative of the function with respect to the corresponding coordinate.*

For polar coordinates we have $\frac{\partial f}{\partial \theta} \neq D_{\mathbf{u}_\theta} f$.

Leaving f out of the formulas in (3) gives us equations of differential operators:

$$D_{\mathbf{u}_r} = D_{\mathbf{v}_r} = \frac{\partial}{\partial r} \qquad \text{and} \qquad D_{\mathbf{u}_\theta} = \frac{1}{r} D_{\mathbf{v}_\theta} = \frac{1}{r}\frac{\partial}{\partial \theta}. \tag{4}$$

Coordinate Vector Fields in Non-orthogonal Coordinates (Optional). If (r, s) are coordinates on E^2, then position is a function of (r, s), that is, $X = X(r, s)$. The same reasoning as above implies that the coordinate vector fields for this coordinate system are $\mathbf{v}_r = \frac{\partial X}{\partial r}$ and $\mathbf{v}_s = \frac{\partial X}{\partial s}$. It is very important to note that these are not necessarily unit vectors, and they are not necessarily orthogonal. In the nicest coordinate systems they are orthogonal, and that is enough to make calculations with them relatively simple. Coordinate vector fields in higher dimensions are similar.

Example 1. Consider E^2 with a Euclidean coordinate system (x, y). On the half of E^2 on which $x > 0$ we define coordinates (r, s) as follows. Given point X with Cartesian coordinates (x, y) with $x > 0$, let

$$r = x \qquad \text{and} \qquad s = y/x.$$

Thus the new coordinates of X are its usual x coordinate and the slope of the line joining X and the origin. Solving for x and y we have

$$x = r \qquad \text{and} \qquad y = rs.$$

The formula for X in terms of (r, s), which you can think of as a map from the rs-plane to the xy-plane, is

$$X = (x, y) = X(r, s) = (r, rs).$$

The coordinate vector fields are then

$$\mathbf{v}_r = \frac{\partial X}{\partial r} = \mathbf{i} + s\mathbf{j} \qquad \text{and} \qquad \mathbf{v}_s = \frac{\partial X}{\partial s} = r\mathbf{j}.$$

Notice that even though $r = x$, the coordinate vector field $\mathbf{v}_r$ is not $\mathbf{i}$. This is because the coordinate vector fields for a coordinate system depend on all of the coordinates, not just the particular coordinate. The vector $\mathbf{v}_r$ points in the direction in which s is constant, so it points along the lines through the origin. It points in the direction r increases, and its length reflects how fast you have to go in order for r to increase by one in one unit of time. Note also that $\mathbf{v}_r$ and $\mathbf{v}_s$ are not orthogonal.

Gradients in non-Euclidean Coordinate Systems

If $f{:}E^2 \to \mathbb{R}$ is differentiable and we express f in Euclidean coordinates (x, y), then the gradient of f is given by $\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$. The generalization of this is the following. At any point in E^2, let $\mathbf{u}_1$ and $\mathbf{u}_2$ be orthogonal unit vectors (an orthonormal basis). Then

$$\nabla f = (D_{\mathbf{u}_1} f)\mathbf{u}_1 + (D_{\mathbf{u}_2} f)\mathbf{u}_2. \tag{5}$$

If you like to think of ∇ as a operator involving vectors and differential operators, then

$$\nabla = \mathbf{u}_1 D_{\mathbf{u}_1} + \mathbf{u}_2 D_{\mathbf{u}_2}, \qquad \text{which generalizes} \qquad \nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y}.$$

These formulas are actually coordinate-free in the sense that the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are not tied to any particular coordinate system. They can be used to compute the gradient of a function in any coordinate system. Formula (5) is particularly easy to use in orthogonal coordinate systems, that is, coordinate systems in which the coordinate vector fields are orthogonal (which happens for polar, cylindrical, and spherical coordinates). In this case, choose $\mathbf{u}_1$ and $\mathbf{u}_2$ to be the unit coordinate vector fields.

Gradients in Polar Coordinates. If we apply formula (5) to polar coordinates with $\mathbf{u}_1 = \mathbf{u}_r$ and $\mathbf{u}_2 = \mathbf{u}_\theta$ and use the derivative formulas in (4), we get

$$\nabla f = (D_{\mathbf{u}_r} f)\mathbf{u}_r + (D_{\mathbf{u}_\theta} f)\mathbf{u}_\theta = \frac{\partial f}{\partial r}\mathbf{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{u}_\theta = \frac{\partial f}{\partial r}\mathbf{v}_r + \frac{1}{r^2}\frac{\partial f}{\partial \theta}\mathbf{v}_\theta. \tag{6}$$

This is much easier than the proof the author of our text has in mind for this formula in Theorem 4.5 of Section 3.4. As an exercise, this method to compute the formula for gradient in spherical coordinates in Theorem 4.6 of Section 3.4.

Gradients in Non-orthogonal Coordinates (Optional). Suppose (r, s) are coordinates on E^2 and we want to determine the formula for ∇f in this coordinate system. In a non-orthogonal coordinate system, applying (5) directly can be messy. The unit vectors $\mathbf{u}_r$ and $\mathbf{u}_s$ aren't orthogonal, and so can't be used for $\mathbf{u}_1$ and $\mathbf{u}_2$. (One of them can be used for $\mathbf{u}_1$, and then $\mathbf{u}_2$ can be $\mathbf{u}_{1\perp}$, but the formula for $\mathbf{u}_2$ as a linear combination of $\mathbf{u}_r$ and $\mathbf{u}_s$ will likely be complicated.) Fortunately, a different approach can be taken.

The coordinate system is well-defined when the coordinate vector fields $\mathbf{v}_r$ and $\mathbf{v}_s$ are independent. If f is a differentiable function, then ∇f will be some linear combination of $\mathbf{v}_r$ and $\mathbf{v}_s$, that is,

$$\nabla f = A\mathbf{v}_r + B\mathbf{v}_s,$$

where the coefficients A and B are some functions of r and s. The procedure for determining A and B is to dot both sides of this equation with $\mathbf{v}_r$ and $\mathbf{v}_s$ to produce two equations in the unknowns A and B.

Example 2. We continue with the coordinates (r, s) on half of E^2 in Example 1. The coordinate vector fields are

$$\mathbf{v}_r = \frac{\partial X}{\partial r} = \mathbf{i} + s\mathbf{j} \qquad \text{and} \qquad \mathbf{v}_s = \frac{\partial X}{\partial s} = r\mathbf{j}.$$

If $f{:}E^2 \to \mathbb{R}$ is differentiable, its gradient can be expressed as a linear combination of $\mathbf{v}_r$ and $\mathbf{v}_s$:

$$\nabla f = A\mathbf{v}_r + B\mathbf{v}_s.$$

Dotting with $\mathbf{v}_r$ and $\mathbf{v}_s$ we get

$$\frac{\partial f}{\partial r} = \nabla f \cdot \mathbf{v}_r = A\mathbf{v}_r{\cdot}\mathbf{v}_r + B\mathbf{v}_s{\cdot}\mathbf{v}_r \qquad \text{and} \qquad \frac{\partial f}{\partial s} = \nabla f \cdot \mathbf{v}_s = A\mathbf{v}_r{\cdot}\mathbf{v}_s + B\mathbf{v}_s{\cdot}\mathbf{v}_s.$$

Note that $\mathbf{v}_r \cdot \mathbf{v}_r = 1 + s^2$, $\mathbf{v}_r \cdot \mathbf{v}_s = rs$, and $\mathbf{v}_s \cdot \mathbf{v}_s = r^2$. Thus we have

$$\frac{\partial f}{\partial r} = A(1 + s^2) + Brs \qquad \text{and} \qquad \frac{\partial f}{\partial s} = Ars + Br^2,$$

which we view as a system of equations with A and B as the unknowns. Solving for A and B we get

$$A = \frac{\partial f}{\partial r} - \frac{s}{r}\frac{\partial f}{\partial s} \qquad \text{and} \qquad B = -\frac{s}{r}\frac{\partial f}{\partial r} + \left(1 + \frac{s^2}{r^2}\right)\frac{\partial f}{\partial s}.$$

This is done easily enough by hand, but more complicated equations may require *Mathematica*. I have added a section to the *Mathematica* notebook Coordinates.nb in the Math 225 folder illustrating this particular computation. Thus we have

$$\nabla f = \left(\frac{\partial f}{\partial r} - \frac{s}{r}\frac{\partial f}{\partial s}\right)\mathbf{v}_r + \left(-\frac{s}{r}\frac{\partial f}{\partial r} + \left(1 + \frac{s^2}{r^2}\right)\frac{\partial f}{\partial s}\right)\mathbf{v}_s.$$

If you apply this process in polar coordinates (it's much easier than in this example because $\mathbf{v}_r$ and $\mathbf{v}_\theta$ are orthogonal), you get

$$\nabla f = \frac{\partial f}{\partial r}\mathbf{v}_r + \frac{1}{r^2}\frac{\partial f}{\partial \theta}\mathbf{v}_\theta,$$

which is the same as (6).

Divergence in non-Euclidean Coordinate Systems

Suppose $\mathbf{F}$ is a vector field on E^2. At every point $\mathbf{F}$ can be written as a linear combination of $\mathbf{i}$ and $\mathbf{j}$, that is,

$$\mathbf{F} = P\mathbf{i} + Q\mathbf{j},$$

where P and Q are functions. In Euclidean coordinates the divergence of $\mathbf{F}$ is given by

$$\operatorname{div}\mathbf{F} = \nabla\cdot\mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y},$$

which can be written as

$$\operatorname{div}\mathbf{F} = \mathbf{i}\cdot\frac{\partial \mathbf{F}}{\partial x} + \mathbf{j}\cdot\frac{\partial \mathbf{F}}{\partial y}.$$

The coordinate-free generalization of this formula is

$$\operatorname{div}\mathbf{F} \;=\; \mathbf{u}_1\cdot D_{\mathbf{u}_1}\mathbf{F} \;+\; \mathbf{u}_2\cdot D_{\mathbf{u}_2}\mathbf{F}, \tag{7}$$

where, as above, $\mathbf{u}_1$ and $\mathbf{u}_2$ form an orthonormal basis. You can still interpret this as $\nabla\cdot\mathbf{F}$ for the generalized definition of ∇ above, *however, the order of operations is important.* The derivatives must be taken before the dot products.

Divergence in Cylindrical Coordinates. Let's apply the three-dimensional version of (7) in cylindrical coordinates. Suppose $\mathbf{F}$ is a vector field on E^3. At every point other than the origin, we can write $\mathbf{F}$ as a linear combination of $\mathbf{u}_r$, $\mathbf{u}_\theta$, and $\mathbf{u}_z = \mathbf{k}$ that is,

$$\mathbf{F} = P\mathbf{u}_r + Q\mathbf{u}_\theta + R\mathbf{u}_z$$

for some functions P, Q, and R. (Note: The coefficient functions P and Q are not the same as the P and Q in Euclidean coordinates, however they are related by change-of-basis formulas, that is, the formulas relating $\mathbf{u}_r$ and $\mathbf{u}_\theta$ to $\mathbf{i}$ and $\mathbf{j}$.) Applying the three-dimensional version of (7) to this, and remembering to use (4), we get

$$\begin{aligned}
\operatorname{div}\mathbf{F} &= \mathbf{u}_r\cdot D_{\mathbf{u}_r}\mathbf{F} \;+\; \mathbf{u}_\theta\cdot D_{\mathbf{u}_\theta}\mathbf{F} \;+\; \mathbf{u}_z\cdot D_{\mathbf{u}_z}\mathbf{F}\\
&= \mathbf{u}_r\cdot\tfrac{\partial \mathbf{F}}{\partial r} \;+\; \mathbf{u}_\theta\cdot\tfrac{1}{r}\tfrac{\partial \mathbf{F}}{\partial \theta} \;+\; \mathbf{u}_z\cdot\tfrac{\partial \mathbf{F}}{\partial z}\\
&= \mathbf{u}_r\cdot\tfrac{\partial}{\partial r}(P\mathbf{u}_r+Q\mathbf{u}_\theta+R\mathbf{u}_z) \;+\; \tfrac{1}{r}\mathbf{u}_\theta\cdot\tfrac{\partial}{\partial \theta}(P\mathbf{u}_r+Q\mathbf{u}_\theta+R\mathbf{u}_z) \;+\; \mathbf{u}_z\cdot\tfrac{\partial}{\partial z}(P\mathbf{u}_r+Q\mathbf{u}_\theta+R\mathbf{u}_z).
\end{aligned}$$

Before going further, we note that $\mathbf{u}_r$ and $\mathbf{u}_\theta$ are not constant vector fields, so they must be differentiated along with the coefficients. (The third unit vector, $\mathbf{u}_z$, is constant, but that is a feature of this particular coordinate system. For the moment we will pretend that it is not constant.) Expanding this without thinking would produce a huge mess: the product rule would result in eighteen terms! Most of them turn out to be zero, and many of these can be determined by thinking about which vectors are orthogonal to each other. First consider

$$\mathbf{u}_\theta\cdot\tfrac{\partial}{\partial \theta}(P\mathbf{u}_r+Q\mathbf{u}_\theta+R\mathbf{u}_z).$$

Applying the product rule to expand $\frac{\partial}{\partial\theta}(P\mathbf{u}_r)$, one of the terms of $\mathbf{u}_\theta\cdot\frac{\partial}{\partial\theta}(P\mathbf{u}_r)$ is $\mathbf{u}_\theta\cdot\frac{\partial P}{\partial\theta}\mathbf{u}_r$. This is zero because $\mathbf{u}_r$ and $\mathbf{u}_\theta$ are orthogonal. One of the terms of $\mathbf{u}_\theta\cdot\frac{\partial}{\partial\theta}(R\mathbf{u}_z)$ is zero for the same reason. If we apply the product rule in $\mathbf{u}_\theta\cdot\frac{\partial}{\partial\theta}(Q\mathbf{u}_\theta)$, we get $\mathbf{u}_\theta\cdot\left(\frac{\partial Q}{\partial\theta}\mathbf{u}_\theta+Q\frac{\partial \mathbf{u}_\theta}{\partial\theta}\right)$. The first part of this is $\frac{\partial Q}{\partial\theta}$ because $\mathbf{u}_\theta$ is a unit vector. For the second part, recall that a vector of constant length is orthogonal to its derivative. Thus, $\mathbf{u}_\theta$ and $\frac{\partial \mathbf{u}_\theta}{\partial\theta}$ are orthogonal, and so the second term for $Q\mathbf{u}_\theta$ is zero.

Putting these observations together, we have

$$\mathbf{u}_\theta\cdot\tfrac{\partial}{\partial\theta}(P\mathbf{u}_r+Q\mathbf{u}_\theta+R\mathbf{u}_z)=P\mathbf{u}_\theta\cdot\tfrac{\partial \mathbf{u}_r}{\partial\theta}+\tfrac{\partial Q}{\partial\theta}+R\mathbf{u}_\theta\cdot\tfrac{\partial \mathbf{u}_z}{\partial\theta}=\mathbf{u}_\theta\cdot\left(P\tfrac{\partial \mathbf{u}_r}{\partial\theta}+\tfrac{\partial Q}{\partial\theta}\mathbf{u}_\theta+R\tfrac{\partial \mathbf{u}_z}{\partial\theta}\right).$$

If we do the same for the other parts of $\operatorname{div}\mathbf{F}$, we get

$$\operatorname{div}\mathbf{F}=\mathbf{u}_r\cdot\left(\tfrac{\partial P}{\partial r}\mathbf{u}_r+Q\tfrac{\partial \mathbf{u}_\theta}{\partial r}+R\tfrac{\partial \mathbf{u}_z}{\partial r}\right)+\tfrac{1}{r}\mathbf{u}_\theta\cdot\left(P\tfrac{\partial \mathbf{u}_r}{\partial\theta}+\tfrac{\partial Q}{\partial\theta}\mathbf{u}_\theta+R\tfrac{\partial \mathbf{u}_z}{\partial\theta}\right)+\mathbf{u}_z\cdot\left(P\tfrac{\partial \mathbf{u}_r}{\partial z}+Q\tfrac{\partial \mathbf{u}_\theta}{\partial z}+\tfrac{\partial R}{\partial r}\mathbf{u}_z\right).$$

Thus, half of the eighteen terms are zero and three of the remaining terms are done (the ones involving the derivatives of P, Q, and R) simply by using the fact that $\mathbf{u}_r$, $\mathbf{u}_\theta$, and $\mathbf{u}_z$ form an orthonormal basis. Moreover, the rest of the computation depends on knowing the dot products of $\mathbf{u}_r$, $\mathbf{u}_\theta$, and $\mathbf{u}_z$ with their derivatives. *This will be true for any orthogonal coordinate system in three dimensions.*

The nine derivatives of $\mathbf{u}_r$, $\mathbf{u}_\theta$, and $\mathbf{u}_z$ with respect to r, θ, and z are particularly simple. From

$$\mathbf{u}_r=\cos\theta\,\mathbf{i}+\sin\theta\,\mathbf{j}\qquad \mathbf{u}_\theta=-\sin\theta\,\mathbf{i}+\cos\theta\,\mathbf{j}\qquad\text{and}\qquad \mathbf{u}_z=\mathbf{k}$$

we have

$$\frac{\partial \mathbf{u}_r}{\partial\theta}=-\sin\theta\,\mathbf{i}+\cos\theta\,\mathbf{j},\qquad\text{and}\qquad \frac{\partial \mathbf{u}_\theta}{\partial\theta}=-\cos\theta\,\mathbf{i}-\sin\theta\,\mathbf{j},$$

and all of the other derivatives are $\mathbf{0}$. It is convenient to express these in terms of the basis $\mathbf{u}_r$, $\mathbf{u}_\theta$, $\mathbf{u}_z$. We have

$$\frac{\partial \mathbf{u}_r}{\partial\theta}=\mathbf{u}_\theta,\qquad\text{and}\qquad \frac{\partial \mathbf{u}_\theta}{\partial\theta}=-\mathbf{u}_r.$$

This is easy to simply observe in this case. For other coordinate systems it may not be quite this easy, but it is simplified when the coordinate vectors are orthogonal—see the handout "Two Very Useful Things from Linear Algebra."

Plugging this information into the last expression for $\operatorname{div}\mathbf{F}$ yields

$$\operatorname{div}\mathbf{F} \;=\; \frac{\partial P}{\partial r} + \frac{1}{r}\left(P + \frac{\partial Q}{\partial \theta}\right) + \frac{\partial R}{\partial z} \;=\; \frac{\partial P}{\partial r} + \frac{1}{r}P + \frac{1}{r}\frac{\partial Q}{\partial \theta} + \frac{\partial R}{\partial z}.$$

You should be sure to check this so you see where everything comes from. Since $\frac{1}{r}\frac{\partial}{\partial r}(rP) = \frac{P}{r} + \frac{\partial P}{\partial r}$, some people like to write this as

$$\operatorname{div}\mathbf{F} \;=\; \frac{1}{r}\frac{\partial}{\partial r}(rP) + \frac{1}{r}\frac{\partial Q}{\partial \theta} + \frac{\partial R}{\partial z}, \quad \text{or even as} \quad \operatorname{div}\mathbf{F} \;=\; \frac{1}{r}\left(\frac{\partial}{\partial r}(rP) + \frac{\partial Q}{\partial \theta} + \frac{\partial}{\partial z}(rR)\right).$$

If this seems messy, compare it with the computation of $\operatorname{div}\mathbf{F}$ in cylindrical coordinates in the text, in the proof of Theorem 4.5, Section3.4. You should try verifying the formula for divergence in spherical coordinates (Theorem 4.6).

A Final Comment about Divergence (Optional). If f is a function and $\mathbf{F}$ and $\mathbf{G}$ are vector fields, all on E^n, the following sum and product rules hold for divergence:

$$\operatorname{div}(\mathbf{F}+\mathbf{G}) = \operatorname{div}\mathbf{F} + \operatorname{div}\mathbf{G} \qquad \text{and} \qquad \operatorname{div}(f\mathbf{F}) = \nabla f \cdot \mathbf{F} + f \operatorname{div}\mathbf{F}.$$

These are easy to verify using Euclidean coordinates, or by using the coordinate-free formulas for gradient and divergence in (5) and (7).

Suppose (r, s) is a coordinate system on E^2 with coordinate vectors $\mathbf{v}_r$ and $\mathbf{v}_s$ and unit coordinate vectors $\mathbf{u}_r$ and $\mathbf{u}_s$. If $\mathbf{F}$ is a vector field, then $\mathbf{F} = P\mathbf{u}_r + Q\mathbf{u}_s$ for some coefficients P and Q. If (r, s) is a Euclidean coordinate system, then the divergence of $\mathbf{F}$ is, of course $\frac{\partial P}{\partial r} + \frac{\partial Q}{\partial s}$.

For a general coordinate system (even non-orthogonal) we have (applying the sum and product rules)

$$\begin{aligned}
\operatorname{div}\mathbf{F} &= \operatorname{div}(P\mathbf{u}_r) + \operatorname{div}(Q\mathbf{u}_s) \\
&= \nabla P \cdot \mathbf{u}_r + P \operatorname{div}\mathbf{u}_r + \nabla Q \cdot \mathbf{u}_s + Q \operatorname{div}\mathbf{u}_s \\
&= D_{\mathbf{u}_r}P + P \operatorname{div}\mathbf{u}_r + D_{\mathbf{u}_s}Q + Q \operatorname{div}\mathbf{u}_s \\
&= \frac{1}{\|\mathbf{v}_r\|}D_{\mathbf{v}_r}P + P \operatorname{div}\mathbf{u}_r + \frac{1}{\|\mathbf{v}_s\|}D_{\mathbf{v}_s}Q + Q \operatorname{div}\mathbf{u}_s \\
&= \frac{1}{\|\mathbf{v}_r\|}\frac{\partial P}{\partial r} + P \operatorname{div}\mathbf{u}_r + \frac{1}{\|\mathbf{v}_s\|}\frac{\partial Q}{\partial s} + Q \operatorname{div}\mathbf{u}_s.
\end{aligned}$$

Things are easier if you abandon the unit vectors and use $\mathbf{v}_r$ and $\mathbf{v}_s$ as a basis. Then we have $\mathbf{F} = p\mathbf{v}_r + q\mathbf{v}_s$. These coefficients are not the same as P and Q, however, $p\,\|\mathbf{v}_r\| = P$ and $q\,\|\mathbf{v}_s\| = Q$. An easy modification of the previous computation yields

$$\operatorname{div}\mathbf{F} = \frac{\partial p}{\partial r} + p \operatorname{div}\mathbf{v}_r + \frac{\partial q}{\partial s} + q \operatorname{div}\mathbf{v}_s.$$

Both of these formulas for divergence show how the extra coefficients and terms in the expression for divergence arise from the coordinate system being non-Euclidean. This formula can also be used to compute an explicit formula for divergence in a non-Euclidean coordinate system, even non-orthogonal coordinates.

Laplacian in non-Euclidean Coordinate Systems

If $f{:}E^2 \to \mathbb{R}$ is twice differentiable, the Laplacian of f is the the divergence of ∇f,

$$\Delta f = \operatorname{div}(\nabla f) = \nabla\cdot(\nabla f).$$

In Euclidean coordinates this is $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$. It's popular to write this as $\nabla\cdot\nabla f = (\nabla\cdot\nabla)f = \nabla^2 f$, and this works since $\nabla\cdot\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ in these coordinates. It's tempting to generalize this as

$$\Delta = \nabla\cdot\nabla = (\mathbf{u}_1 D_{\mathbf{u}_1} + \mathbf{u}_2 D_{\mathbf{u}_2})\cdot(\mathbf{u}_1 D_{\mathbf{u}_1} + \mathbf{u}_2 D_{\mathbf{u}_2}) \stackrel{?}{=} D^2_{\mathbf{u}_1} + D^2_{\mathbf{u}_2},$$

but this overlooks a subtlety. The last equality is not true unless $\mathbf{u}_1$ and $\mathbf{u}_2$ are constant vector fields, which of course they aren't when dealing with non-Euclidean coordinate systems. In expanding the dot product, the derivative operators $D_{\mathbf{u}_1}$ and $D_{\mathbf{u}_2}$ in the first factor have to be applied to the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ in the second factor before the dot product is taken.[1] Fortunately, there is an easier way. Assuming we have formulas for gradient and divergence in a coordinate system, we can simply use them to get the formula for Laplacian.

Laplacian in Polar Coordinates. In polar coordinates we have

$$\nabla f = \frac{\partial f}{\partial r}\mathbf{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{u}_\theta \qquad \text{and} \qquad \operatorname{div}(P\mathbf{u}_r + Q\mathbf{u}_\theta) = \frac{1}{r}\frac{\partial}{\partial r}(rP) + \frac{1}{r}\frac{\partial Q}{\partial \theta}.$$

Thus,

$$\Delta f = \operatorname{div}\left(\frac{\partial f}{\partial r}\mathbf{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{u}_\theta\right) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r}\frac{\partial}{\partial \theta}\left(\frac{1}{r}\frac{\partial f}{\partial \theta}\right) = \frac{1}{r}\frac{\partial f}{\partial r} + \frac{\partial^2 f}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}.$$

This formula is in Exercise 31 of Section 2.5, but there it is expressed without the function f. You should try deriving the formula for Laplacian in spherical coordinates (Exercise 33, Section 2.5).

Curl in non-Euclidean Coordinate Systems

In E^2 we have $\operatorname{curl}\mathbf{F} = \nabla_\perp\cdot\mathbf{F}$ and in E^3 we have $\operatorname{curl}\mathbf{F} = \nabla\times\mathbf{F}$.

[1] The resulting formula is somewhat interesting: $\nabla\cdot\nabla = D^2_{\mathbf{u}_1} + D^2_{\mathbf{u}_2} + (\operatorname{div}\mathbf{u}_1)D_{\mathbf{u}_1} + (\operatorname{div}\mathbf{u}_2)D_{\mathbf{u}_2}$.

Two Dimensions. If $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$ on E^2, then

$$\operatorname{curl}\mathbf{F} = \nabla_{\perp}\cdot\mathbf{F} = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y}\right)_{\perp}\cdot\mathbf{F} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \det\begin{pmatrix}\partial/\partial x & \partial/\partial y\\ P & Q\end{pmatrix}.$$

This can also be written as

$$\operatorname{curl}\mathbf{F} \;=\; \nabla_{\perp}\cdot\mathbf{F} \;=\; \mathbf{i}_{\perp}\cdot\frac{\partial\mathbf{F}}{\partial x} \;+\; \mathbf{j}_{\perp}\cdot\frac{\partial\mathbf{F}}{\partial y}.$$

The generalization of this for orthonormal basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ is

$$\operatorname{curl}\mathbf{F} \;=\; (\mathbf{u}_1)_{\perp}\cdot D_{\mathbf{u}_1}\mathbf{F} \;+\; (\mathbf{u}_2)_{\perp}\cdot D_{\mathbf{u}_2}\mathbf{F}. \tag{8}$$

Once again, this can be interpreted as $\nabla_{\perp}\cdot\mathbf{F}$ as long as you are careful to take the derivatives before the dot products.

The orthonormal basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are said to be *positively-oriented* if $\mathbf{u}_2 = (\mathbf{u}_1)_{\perp}$. Note that the standard basis vectors $\mathbf{i}$ and $\mathbf{j}$ have this property. (The order is important—the second basis vector is the perp of the first.) With this assumption, and stretching the meaning of determinant a bit, $\operatorname{curl}\mathbf{F}$ can be written as

$$\operatorname{curl}\mathbf{F} \;=\; \nabla_{\perp}\cdot\mathbf{F} \;=\; \mathbf{u}_2\cdot D_{\mathbf{u}_1}\mathbf{F} \;-\; \mathbf{u}_1\cdot D_{\mathbf{u}_2}\mathbf{F} \;=\; \det\begin{pmatrix}D_{\mathbf{u}_1}\mathbf{F} & D_{\mathbf{u}_2}\mathbf{F}\\ \mathbf{u}_1 & \mathbf{u}_2\end{pmatrix}.$$

Note that the entries of this matrix are vectors, and the product you use when expanding it is dot product. I'm not convinced that this is a useful formula.

Curl in Polar Coordinates. Let's use (8) to compute $\operatorname{curl}\mathbf{F}$ in polar coordinates. The computation is very similar to that of $\operatorname{div}\mathbf{F}$ above. Assume $\mathbf{F} = P\mathbf{u}_r + Q\mathbf{u}_\theta$. We have

$$\begin{aligned}\operatorname{curl}\mathbf{F} \;&=\; (\mathbf{u}_r)_{\perp}\cdot D_{\mathbf{u}_r}\mathbf{F} \;+\; (\mathbf{u}_\theta)_{\perp}\cdot D_{\mathbf{u}_\theta}\mathbf{F} \;=\; \mathbf{u}_\theta\cdot\frac{\partial\mathbf{F}}{\partial r} \;-\; \mathbf{u}_r\cdot\frac{1}{r}\frac{\partial\mathbf{F}}{\partial\theta}\\ &=\; \mathbf{u}_\theta\cdot\frac{\partial}{\partial r}(P\mathbf{u}_r + Q\mathbf{u}_\theta) \;-\; \mathbf{u}_r\cdot\frac{1}{r}\frac{\partial}{\partial\theta}(P\mathbf{u}_r + Q\mathbf{u}_\theta)\\ &=\; \frac{\partial Q}{\partial r} \;-\; \frac{1}{r}\left(\frac{\partial P}{\partial\theta} - Q\right) \;=\; \frac{1}{r}\left(\frac{\partial}{\partial r}(rQ) - \frac{\partial P}{\partial\theta}\right) \;=\; \frac{1}{r}\det\begin{pmatrix}\partial/\partial r & \partial/\partial\theta\\ P & rQ\end{pmatrix}.\end{aligned}$$

Three Dimensions. If $\mathbf{F}$ is a vector field on E^3, then

$$\operatorname{curl}\mathbf{F} = \nabla\times\mathbf{F} = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\times\mathbf{F} = \mathbf{i}\times\frac{\partial\mathbf{F}}{\partial x} + \mathbf{j}\times\frac{\partial\mathbf{F}}{\partial y} + \mathbf{k}\times\frac{\partial\mathbf{F}}{\partial z}.$$

By now you can probably guess what the generalization of this is. If $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ form an orthonormal basis, then

$$\begin{aligned}\operatorname{curl}\mathbf{F} &= \nabla\times\mathbf{F} = (\mathbf{u}_1 D_{\mathbf{u}_1} + \mathbf{u}_2 D_{\mathbf{u}_2} + \mathbf{u}_3 D_{\mathbf{u}_3})\times\mathbf{F}\\ &=\; \mathbf{u}_1\times D_{\mathbf{u}_1}\mathbf{F} \;+\; \mathbf{u}_2\times D_{\mathbf{u}_2}\mathbf{F} \;+\; \mathbf{u}_3\times D_{\mathbf{u}_3}\mathbf{F}.\end{aligned}$$

While this can be massaged into a determinant formula, it takes a bit of work, isn't very enlightening, and stretches the notion of determinant even more.

You should try deriving the formulas for curl in cylindrical and spherical coordinates.

Coordinate Vector Fields, Gradient, Divergence, Laplacian, and Curl in Other Settings

Let S be a surface in E^3, for example, a sphere or torus. The ideas presented above can be used to define these concepts on S. We have seen examples of coordinates on surfaces, and gradient, divergence, etc., can be expressed in those coordinates.

Robert L. Foote, October 2005
Revised October 2014