'Coordinate Systems and Examples of the Chain Rule

Alex Nita

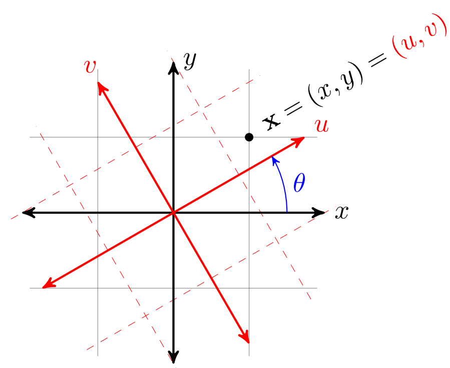
Abstract

One of the reasons the chain rule is so important is that we often want to change coordinates in order to make difficult problems easier by exploiting internal symmetries or other nice properties that are hidden in the Cartesian coordinate system. We will see how this works for example when trying to solve first order linear partial differential equations or when working with differential equations on circles, spheres and cylinders. It is one of the achievements of vector calculus that it has a develped language for describing all of these different aspects in a unified way, in particular using coordinate changes and differentiating functions under these coordinate changes to provide simplified expressions. The key idea is the following: if $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is a change-of-coordinates function and $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, then instead of working with $Df(\mathbf{x})$, which may be difficult in practice, we may work with $D(f \circ \varphi)(\mathbf{x}) = Df(\varphi(\mathbf{x}))D\varphi(\mathbf{x})$.

1 Different Coordinate Systems in $\mathbb{R}^2$ and $\mathbb{R}^3$

Example 1.1 *Let's start with something simple, like rotating our coordinates through an angle of θ. In a previous example, in the section on differentiability, we showed that a rotation through θ is a linear function $R : \mathbb{R}^2 \to \mathbb{R}^2$ which operates on vectors (x, y), written here as column vectors, by left multiplying them by the rotation matrix R,*

$$\begin{pmatrix} u \\ v \end{pmatrix} = R \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} \tag{1.1}$$



In terms of the component functions *of R, we have*

$$u = R_1(x, y) = x\cos\theta - y\sin\theta$$
$$v = R_2(x, y) = x\sin\theta + y\cos\theta$$

We have thus introduced new **rotated coordinates** (u, v) *in* $\mathbb{R}^2$. ■

Example 1.2 *There are other coordinate systems in $\mathbb{R}^2$ besides rotated Cartesian coordinates, but which are still "linear" in a sense. Consider for example two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$ which are not scalar multiples of each other (i.e. they are not collinear). Then I claim that* any *vector $\mathbf{x}$ in $\mathbb{R}^2$ can be described as a scalar multiple of $\mathbf{u}$ plus a scalar multiple of $\mathbf{v}$ (hence they describe a* coordinate system, *telling you how to find $\mathbf{x}$ by going some way in the direction of $\mathbf{u}$ then some way in the direction of $\mathbf{v}$, much like in regular Cartesian coordinates we go some way in the direction of $\mathbf{i}$ and some way in the direction of $\mathbf{j}$). Let's see how this works. Let*

$$\mathbf{u} = (a, b) \qquad \textit{and} \qquad \mathbf{v} = (c, d)$$

and suppose they are not scalar multiples of each other (i.e. $(c, d) \neq k(a, b) = (ka, kb)$ for any k). Then consider any $\mathbf{x} = (x, y)$. To say that $\mathbf{x} = k_1\mathbf{u} + k_2\mathbf{v}$ is to say

$$(x, y) = k_1(a, b) + k_2(c, d) = \big(k_1a + k_2c,\ k_1b + k_2d\big)$$

which, if we want to discover what k_1 and k_2 are, means we have to solve a system of two equations in two unknowns,

$$\begin{aligned} x &= ak_1 + bk_2 \\ y &= bk_1 + dk_2 \end{aligned}$$

(Recall that here a, b, c, d and x, y are all known, only k_1 and k_2 are unkown!) OK, well, we know how to solve such a system, namely add $-c/a$ times the first equation to the second equation,

$$\begin{cases} x = ak_1 + bk_2 \\ y = ck_1 + dk_2 \end{cases} \implies \begin{cases} x = ak_1 + bk_2 \\ -\frac{c}{a}x + y = 0k_1 + \left(-\frac{bc}{a} + d\right)k_2 \end{cases}$$

and then solve the second equation for k_2,

$$k_2 = \frac{ay - cx}{a} \cdot \frac{a}{ad - bc} = \frac{ay - cx}{ad - bc}$$

Finally, substitute k_2 back into either equation and solve the result for k_1. For example, substitute it into the first equation:

$$x = ak_1 + bk_2 = ak_1 + b\left(\frac{ay - cx}{ad - bc}\right) \implies k_1 = \frac{x - b\left(\frac{ay-cx}{ad-bc}\right)}{a} = \frac{dx - by}{ad - bc}$$

Thus,

$$\mathbf{x} = k_1\mathbf{u} + k_2\mathbf{v} = \left(\frac{dx - by}{ad - bc}\right)\mathbf{u} + \left(\frac{ay - cx}{ad - bc}\right)\mathbf{v} \tag{1.2}$$

If we want to define a function *$\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ changing $\mathbf{ij}$-coordinates to $\mathbf{uv}$-coordinates, it is the function sending (x, y) to (k_1, k_2), namely*

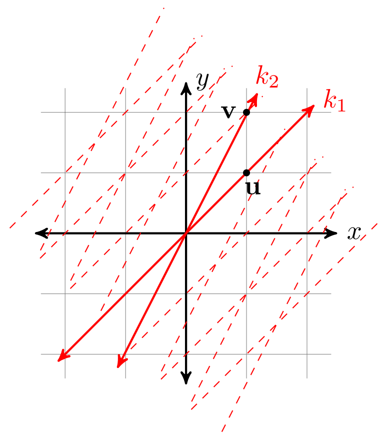
$$\begin{aligned} &\varphi : \mathbb{R}^2 \to \mathbb{R}^2 \\ &(k_1, k_2) = \varphi(x, y) = \left(\frac{dx - by}{ad - bc}, \frac{ay - cx}{ad - bc}\right) \end{aligned} \tag{1.3}$$

We can express this in matrix notation (and this explains the "linearity" of the $\mathbf{uv}$-coordinates):

$$\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \varphi\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{1.4}$$

■

Example 1.3 *Let's try this out with some specific vectors. Let* $\mathbf{u} = (1,1)$ *and* $\mathbf{v} = (1,2)$ *and let* $\mathbf{x} = (3,5)$.



Let's find k_1 *and* k_2 *such that* $\mathbf{x} = k_1\mathbf{u} + k_2\mathbf{v}$, *and check that the above formula works. Well,*

$$k_1 = \frac{dx - by}{ad - bc} = \frac{2 \cdot 3 - 1 \cdot 5}{1 \cdot 2 - 1 \cdot 1} = 1$$
$$k_2 = \frac{ay - cx}{ad - bc} = \frac{1 \cdot 5 - 1 \cdot 3}{1 \cdot 2 - 1 \cdot 1} = 2$$

and indeed

$$k_1\mathbf{u} + k_2\mathbf{v} = 1(1,1) + 2(1,2) = (1,1) + (2,4) = (1+2, 1+4) = (3,5) = \mathbf{x}$$

And we can do this for any $\mathbf{x}$ *in* $\mathbb{R}^2$. ■

There are of course other coordinate systems, and the most common are *polar*, *cylindrical* and *spherical*. Let us discuss these in turn.

Example 1.4 **Polar coordinates** *are used in* $\mathbb{R}^2$, *and specify any point* $\mathbf{x}$ *other than the origin, given in Cartesian coordinates by* $\mathbf{x} = (x, y)$, *by giving the* length r *of* $\mathbf{x}$ *and the* angle *which it makes with the* x*-axis,*

$$r = \|\mathbf{x}\| = \sqrt{x^2 + y^2} \qquad \text{and} \qquad \theta = \arctan\left(-\frac{y}{x}\right) \tag{1.5}$$

Notice that the domain of arctan *is naturally* $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, *so this set of coordinates only works on the right half-plane* $\{(x, y) \mid x > 0\}$. *We can adjoin the* y*-axis and the left half of the plane, but there we have to define* θ *differently. We leave the tedious details to the reader. We then get the composite function*

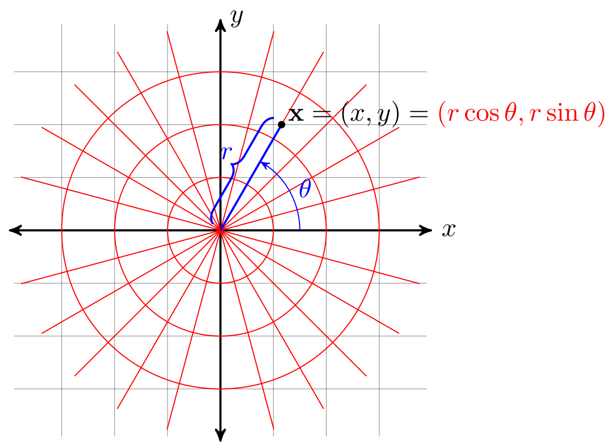
$$\begin{aligned} &\varphi : \mathbb{R}^2 \to \mathbb{R}^2 \\ &(r, \theta) = \varphi(x, y) = \left(\sqrt{x^2 + y^2},\ \arctan\left(-\frac{y}{x}\right)\right) \end{aligned} \tag{1.6}$$

The thing to note is that we can go back and forth between Cartesian and polar coordinates. The reverse direction, from (r, θ) *to* (x, y) *is*

$$\begin{aligned} &\psi = \varphi^{-1} : \mathbb{R}^2 \to \mathbb{R}^2 \\ &(x, y) = \psi(r, \theta) = \big(r\cos\theta,\ r\sin\theta\big) \end{aligned} \tag{1.7}$$

We usually see this in the form,

$$x = r\cos\theta \qquad \textit{and} \qquad y = r\sin\theta \tag{1.8}$$



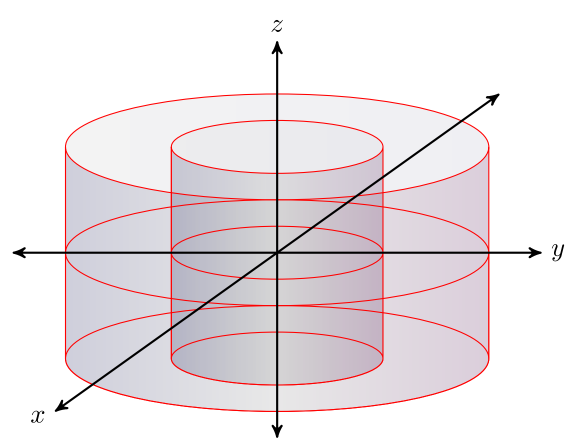
Example 1.5 **Cylindrical coordinates** *are like polar coordinates, but we include a third dimension as well, which we call z (we may as well identify it with the z-coordinate in Cartesian coordinates, though we can just as well identify it with any line through the origin). Thus, our change-of-coordinates functions changing Cartesian into cylindrical coordinates is*

$$\begin{aligned} &\varphi : \mathbb{R}^3 \to \mathbb{R}^3 \\ &(r, \theta, z) = \varphi(x, y, z) = \left(\sqrt{x^2 + y^2},\ \arctan\left(-\frac{y}{x}\right),\ z \right) \end{aligned} \tag{1.9}$$

and its inverse is the map

$$\begin{aligned} &\psi = \varphi^{-1} : \mathbb{R}^3 \to \mathbb{R}^3 \\ &(x, y, z) = \psi(r, \theta, z) = \big(r\cos\theta,\ r\sin\theta,\ z\big) \end{aligned} \tag{1.10}$$

■



Example 1.6 **Spherical coordinates** *are used when working with a system having inherent spherical symmetry, for example the gravitational or the electric field surrounding a point particle.*

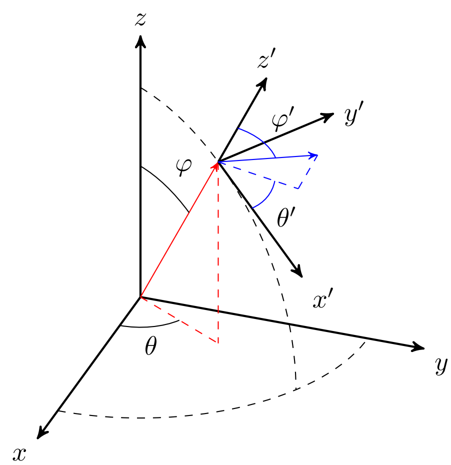


Figure 1.1: Spherical Coordinates

The basic idea behind spherical coordinates is that a point $\mathbf{x} = (x, y, z)$ *can be entirely determined not only by the coordinates* x,y *and* z *in the coordinate directions* $\mathbf{i}$, $\mathbf{j}$ *and* $\mathbf{k}$, *but by it's* length

$$\rho = \|\mathbf{x}\|$$

and two angles, one between $\mathbf{x}$ *and* $\mathbf{k}$ *(the* z*-axis),*

$$\varphi = \arccos\left(\frac{\mathbf{x} \cdot \mathbf{k}}{\|\mathbf{x}\|\|\mathbf{k}\|}\right) = \arccos\left(\frac{z}{\rho}\right)$$

and one between the projection of $\mathbf{x}$ *onto the* x, y*-plane and the* x*-axis, i.e. between* $(x, y, 0)$ *and* $\mathbf{i} = (1, 0, 0)$,

$$\theta = \arccos\left(\frac{(x, y, 0) \cdot (1, 0, 0)}{\|(x, y, 0)\|\|(1, 0, 0)\|}\right) = \arccos\left(\frac{x}{\sqrt{x^2 + y^2}}\right)$$

This could also be described as the arctangent of y/x, *if we draw a right triangle with legs* x *and* y *and hypotenuse* $\sqrt{x^2 + y^2}$

$$\theta = \arctan\left(\frac{y}{x}\right)$$

and there is a similar expression involving arcsine. This gives us a function from $\mathbb{R}^3$ *to* $\mathbb{R}^3$, *converting Cartesian coordinates into spherical coordinates:*

$$\mathcal{C} : \mathbb{R}^3 \to \mathbb{R}^3$$
$$(\rho, \theta, \varphi) = \mathcal{C}(x, y, z) = \left(\|\mathbf{x}\|,\ \arccos\left(\frac{z}{\|\mathbf{x}\|}\right),\ \arctan\left(\frac{y}{x}\right)\right) \tag{1.11}$$

In components this is

$$\begin{aligned}
\rho &= \|\mathbf{x}\| \\
\theta &= \arccos\Big(\frac{z}{\|\mathbf{x}\|}\Big) \\
\varphi &= \arctan\Big(\frac{y}{x}\Big) = \arccos\Big(\frac{x}{\sqrt{x^2+y^2}}\Big) = \arcsin\Big(\frac{y}{\sqrt{x^2+y^2}}\Big)
\end{aligned} \tag{1.12}$$

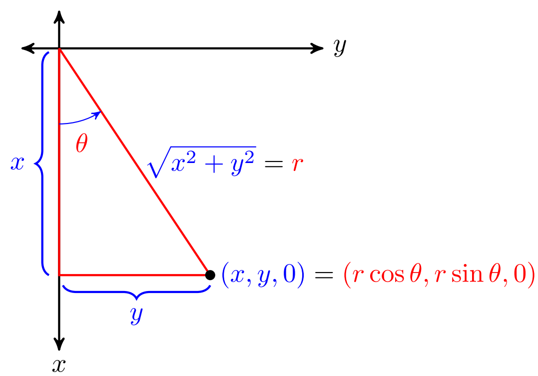
The reverse transformation, $\mathcal{S} = \mathcal{C}^{-1}$, converting spherical to Cartesian coordinates, is

$$\begin{aligned}
&\mathcal{S} = \mathcal{C}^{-1} : \mathbb{R}^3 \to \mathbb{R}^3 \\
&(x, y, z) = \mathcal{S}(\rho, \theta, \varphi) = \Big(\rho \sin\varphi \cos\theta,\ \rho \sin\varphi \sin\theta,\ \rho \cos\varphi\Big)
\end{aligned} \tag{1.13}$$

which, in components is

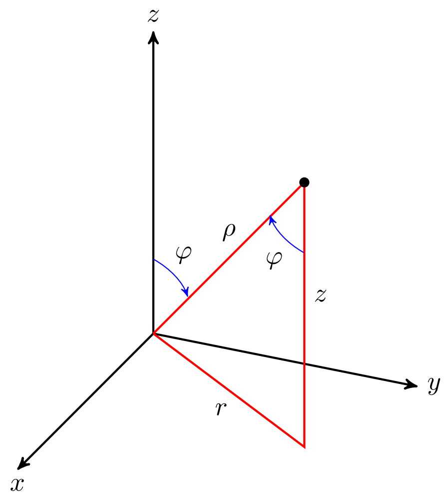
$$\begin{aligned}
x &= \rho \sin\varphi \cos\theta \\
y &= \rho \sin\varphi \sin\theta \\
z &= \rho \cos\varphi
\end{aligned} \tag{1.14}$$

Refer to Figure 1.6 for the derivation to follow: First, note that in that figure, the projection of $\mathbf{x} = (x, y, z)$ onto the xy-plane is $(x, y, 0)$, so θ is easy to compute using the triangle



Indeed, θ is easily seen to be the arctangent of y/x, or the arccosine of $x/\sqrt{x^2+y^2}$, or the arcsine of $y/\sqrt{x^2+y^2}$. In fact, in the xy-plane we have the standard polar coordinates. Hence we also have $x = r\cos\theta$ and $y = r\sin\theta$.

Now let's look at the 3D picture. The vertical triangle has horizontal leg $r = \sqrt{x^2+y^2}$, *vertical leg z, and hypotenuse $\rho = \sqrt{r^2+z^2} = \sqrt{x^2+y^2+z^2} = \|\mathbf{x}\|$,*



From this picture we can see that $\frac{r}{\rho} = \sin\varphi$, so that $r = \rho\sin\varphi$. We can substitute this expression into our earlier expressions for x and y in polar coordinates:

$$x = r\cos\theta = \rho\sin\varphi\cos\theta$$
$$y = r\sin\theta = \rho\sin\varphi\sin\theta$$

Finally, $\frac{z}{\rho} = \cos\varphi$, so we also have

$$z = \rho\cos\varphi$$

■

2 New Coordinates and the Chain Rule

What are coordinates good for? The answer is they are basically the most important thing in vector calculus, because they are precisely what simplifies complicated problems. The only difficulty in switching to more convenient coordinates is that you have to then translate the whole problem you're looking at into these new coordinates, and there may well be need to use the chain rule. The reason is most interesting problems in physics and engineering are equations involving partial derivatives, that is partial differential equations. These equations normally have physical interpretations and are derived from observations and experimentation. Once handed to us, however, we must treat the problems mathematically, and our purpose here is to elucidate some of the basic methods of this mathematical treatment. Let us look at some examples.

Example 2.1 *Suppose you are asked to solve the following first order linear* **partial differential equation***,*

$$3u_x + 4u_y = 0 \qquad \text{or, in other notation,} \qquad 3\frac{\partial u}{\partial x} + 4\frac{\partial u}{\partial y} = 0$$

A **solution** *of this equation is a function $u : \mathbb{R}^2 \to \mathbb{R}$, $z = u(x, y)$, whose partial derivatives satisfy this equation at all points (x, y) in the plane. How are we to solve such a problem?*

Well, the first thing to notice is that what we have here is in fact a matrix (or dot) product on the left hand side, namely

$$3u_x + 4u_y = \begin{pmatrix} u_x & u_y \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = Du \cdot \mathbf{v}$$

where $\mathbf{v} = (3,4)$ *is thought of as a column vector here. But we also know that this product is precisely the* directional derivative *of* u *in the direction of* $\mathbf{v}$*,*

$$D_{\mathbf{v}}u = Du \cdot \mathbf{v} = 3u_x + 4u_v$$

Thus, our original equation boils down to solving the directional derivative equation

$$D_{\mathbf{v}}u = 0$$

Well, what does it say about u *that it's directional derivative in the direction of* $\mathbf{v} = (3,4)$ *is* 0*? It means* u *is not changing in that direction, at any point in* $\mathbb{R}^2$*! Pick a line parallel to* $\mathbf{v}$*, that is having slope* $4/3$*,*

$$y = \frac{4}{3}x + b \qquad \text{or} \qquad 4x - 3y = -3b$$

where b *is arbitrary. Then, on this line,* u *will be constant. The only change in* u *will occur if we move from one line to another, that is in a direction perpendicular to* $\mathbf{v}$ *(for we note, of course, that the equation of the line can be expressed as a dot product, the vector* $(4,-3)$ *with the vector* (x,y)*, and the vector* $(4,-3)$ *is orthogonal to* $\mathbf{v}$*!). Thus, any differentiable function* $f : \mathbb{R} \to \mathbb{R}$ *taking* $\eta = 4x - 3y$ *as input will work for us, so a general solution is*

$$\boxed{u = f \circ g, \quad \text{where} g : \mathbb{R}^2 \to \mathbb{R},\ g(x,y) = 4x - 3y,\ \text{and } f : \mathbb{R} \to \mathbb{R} \text{ is arbitrary}}$$

If we chose coordinates as in Example 1.2 in such a way that one of the coordinate vectors were $\mathbf{v}$*, then the fact that* $D_{\mathbf{v}}u = 0$ *would mean that it's partial derivative with respect to that new coordinate would be zero, always! Consequently, we would only have to worry about the other partial of* u*, in the other coordinate direction. Let's call the other coordinate* $\mathbf{w}$*, which we may as well pick to be orthogonal to* $\mathbf{v}$*. Thus, if say* ξ *and* η *are our new coordinates, meaning that*

$$\xi\mathbf{v} + \eta\mathbf{w} = (x,y)$$

for all (x,y) *in* $\mathbb{R}^2$*, then our original partial differential equation would simplify to*

$$3u_x + 4u_y = 0 \quad \Longrightarrow \quad u_\eta = 0$$

Let us verify that this is indeed a solution: Let $u(x,y) = (f \circ g)(x,y) = f(4x - 3y)$*. Then by the chain rule we have*

$$\begin{aligned} 3u_x + 4u_y &= 3\frac{\partial}{\partial x}f(4x-3y) + 4\frac{\partial}{\partial y}f(4x-3y) \\ &= 3f'(4x-3y)\cdot 4 + 4f'(4x-3y)\cdot(-3) \\ &= (12-12)f'(4x-3y) \\ &= 0 \end{aligned}$$

For example, we could choose $f(t) = e^t$*, or* $f(t) = \sin t$*, or* $f(t) = t^3 + 5$*, or anything we want as long as it's differentiable on* $\mathbb{R}$*! In these cases,* $u(x,y) = e^{4x-3y}$*, or* $u(x,y) = \sin(4x-3y)$*, or* $u(x,y) = (4x-3y)^3 + 5$ *are all solutions.*

There is another way to view this problem which highlights the change of coordinates. This still involves the observation that u *is constant along the lines* $4x - 3y = c$*. The idea, then, is to pick coordinates, one parallel to* $\mathbf{v}$*, the other perpendicular to* $\mathbf{v}$*, for then the partial in one direction will be identically zero and the result will be a simpler partial differential equation. Let us explain.*

Define new coordinates

$$\xi = 3x + 4y$$
$$\eta = 4x - 3y$$

Note that ξ *is in the direction of* $\mathbf{v}$*, and* η *is perpendicular to it. If we solve this system for* x *and* y *this will be clearer:*

$$x = \frac{3}{25}\xi + \frac{4}{25}\eta$$
$$y = \frac{4}{25}\xi - \frac{3}{25}\eta$$

and therefore

$$(x, y) = \left(\frac{3}{25}\xi + \frac{4}{25}\eta,\ \frac{4}{25}\xi - \frac{3}{25}\eta\right) = \xi\frac{1}{25}(3, 4) + \eta\frac{1}{24}(4, -3)$$

so (x, y) *is expressed as a multiple of* $\mathbf{v}$ *plus a multiple of* $\mathbf{w}$*, the vector orthogonal to* $\mathbf{v}$*. What happens when we consider* $u(\xi, \eta)$*, that is the expression of* u *in these new coordinates? Well, let's find out by computing some partials:*

$$u_x = u_\xi \xi_x + u_\eta \eta_x = 3u_\xi + 4u_\eta$$
$$u_y = u_\xi \xi_y + u_\eta \eta_y = 4u_\xi - 3u_\eta$$

and consequently,

$$0 = 3u_x + 4u_y = 3(3u_\xi + 4u_\eta) + 4(4u_\xi - 3u_\eta) = 25u_\xi + (12 - 12)u_\eta = 25u_\xi$$

Therefore, dividing by 25 *gives*

$$u_\xi = 0$$

a much simpler *partial differential equation than* $3u_x + 4u_y = 0$*! Indeed, we can solve this by integrating with respect to* ξ*. When we do, since* u *is a function of* ξ and η*, we will get an arbitrary differentiable function of* η*,*

$$\boxed{u(\xi, \eta) = f(\eta)}$$

Converting back to (x, y) *coordinates now we see that*

$$\boxed{u(x, y) = f(4x - 3y)}$$

■

Exercise 2.2 *Show that any* **homogeneous linear partial differential equation with constant coefficients** *(the technical term for these things),*

$$au_x + bu_y = 0$$

has a general solution given by

$$u(x, y) = f(bx - ay)$$

for any *twice differentiable function* $f : \mathbb{R} \to \mathbb{R}$*.* ■

Exercise 2.3 *Show that any* **inhomogeneous linear partial differential equation with constant coefficients** *(the technical term for these things),*

$$au_x + bu_y = c$$

(where c is any nonzero constant) has a general solution given by

$$u(x, y) = f(bx - ay)e^{-cx/(a^2+y^2)}$$

for any *twice differentiable function $f : \mathbb{R} \to \mathbb{R}$.* [Hint: An ordinary differential equation of the form $ay' + by = 0$ for y a function of x (which we hope to discover, for that would be 'solving' the differential equation), and constants a and b can be solved by noticing that we can rewrite it as $y'/y = -b/a$, then noting that $y'/y = (\ln y)'$ via the chain rule, so that in fact we have $(\ln y)' = -b/a$. Integrating with respect to x gives $\ln y = -(b/a)x + c$ for some arbitrary constant of integration c. Taking the exponential of both sides gives $y = e^{-(b/a)x+c} = Ce^{-(b/a)x}$, where $C = e^c$. Thus, we have solved our differential equation, up to a constant C.] ■

Example 2.4 *Consider the* **Laplace operator** *on $\mathbb{R}^2$,*

$$\boxed{\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}}$$

This operator takes twice differentiable functions $f : \mathbb{R}^2 \to \mathbb{R}$ to at least continuous functions from $\mathbb{R}^2$ to $\mathbb{R}$,

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} : \mathbb{R}^2 \to \mathbb{R}$$

This is the expression of Δ in standard Cartesian coordinates*, but what if we wanted to exploit some internal symmetry of a problem. Suppose that we were trying to solve a* second order differential equation*, called the* **wave equation** *(which obviously models wave propagation through some medium),*

$$\frac{\partial^2 u}{\partial t^2} = \Delta u$$

(Here $u : \mathbb{R}^3 \to \mathbb{R}$ is given also as a function of time*, $u(t, x, y)$, and the z-value of u at a point (x, y) in $\mathbb{R}^2$ at time t represents the* height *of the wave.) Suppose further that we know the wave has some circular symmetry, that is the wave propagates outward symmetrically about some point. Then it would obviously be more convenient to treat this problem, that is to look for solutions to $u_{tt} = \Delta u$, if we convert to polar coordinates. My claim is that the* **Laplace operator in polar coordinates** *takes the expression*

$$\boxed{\Delta = \frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}}$$

Let's prove this by starting with the right-hand side and then showing it's equal to the left-hand side. Well, this means using the polar coordinates of Example 1.4,

$$x = r\cos\theta \qquad \textit{and} \qquad y = r\sin\theta$$

and the chain rule on any twice-differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$, since what we want to compute is the partials of $f(r, \theta)$, which means the partials of the composition $f \circ \psi : \mathbb{R}^2 \to \mathbb{R}$ where

$$\psi = \varphi^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$$
$$(x, y) = \psi(r, \theta) = \big(r\cos\theta,\ r\sin\theta\big)$$

is the change-of-coordinate function from polar to Cartesian. Well, the chain rule in components is

$$\begin{aligned} f_r(r,\theta) &= f_x(\psi(r,\theta))x_r(r,\theta) + f_y(\psi(r,\theta))y_r(r,\theta) \\ f_\theta(r,\theta) &= f_x(\psi(r,\theta))x_\theta(r,\theta) + f_y(\psi(r,\theta))y_\theta(r,\theta) \end{aligned}$$

which we'll abbreviate, in the interest of space, to

$$\begin{aligned} f_r &= f_x x_r + f_y y_r \\ f_\theta &= f_x x_\theta + f_y y_\theta \end{aligned}$$

thus avoiding mention of where we evaluate each (but as you have noted in your WebWork and homework, where you evaluate e.g. f_x is different from where you evaluate y_r!). Then, since

$$x_r = \cos\theta, \quad y_r = \sin\theta, \quad x_\theta = -r\sin\theta, \quad y_\theta = r\cos\theta$$

we have

$$\boxed{\begin{aligned} &f_r = f_x\cos\theta + f_y\sin\theta \\ &f_\theta = -f_x r\sin\theta + f_y r\cos\theta \end{aligned}}$$

Let's now compute the second partials. This will involve using the chain rule again on f_x and f_y, for example $(f_x)_r = (f_x)_x x_r + (f_x)_y y_r = f_{xx}x_r + f_{xy}y_r = f_{xx}\cos\theta + f_{xy}\sin\theta$. So,

$$\begin{aligned} f_{rr} &= (f_r)_r \\ &= \Big(f_x\cos\theta + f_y\sin\theta\Big)_r \\ &= \big(f_x\big)_r\cos\theta + \big(f_y\big)_r\sin\theta \\ &= \Big(f_{xx}x_r + f_{xy}y_r\Big)\cos\theta + \Big(f_{yx}x_r + f_{yy}y_r\Big)\sin\theta \\ &= f_{xx}\cos^2\theta + 2f_{xy}\sin\theta\cos\theta + f_{yy}\sin^2\theta \end{aligned}$$

and

$$\begin{aligned} f_{\theta\theta} &= (f_\theta)_\theta \\ &= \Big(-f_x r\sin\theta + f_y r\cos\theta\Big)_\theta \\ &= -r\big(f_x\sin\theta\big)_\theta + r\big(f_y\cos\theta\big)_\theta \\ &= -r\Big(\big(f_x\big)_\theta\sin\theta + f_x\big(\sin\theta\big)_\theta\Big) + r\Big(\big(f_y\big)_\theta\cos\theta + f_y\big(\cos\theta\big)_\theta\Big) \\ &= -r\Big(\big(f_{xx}x_\theta + f_{xy}y_\theta\big)\sin\theta + f_x\cos\theta\Big) + r\Big(\big(f_{yx}x_\theta + f_{yy}y_\theta\big)\cos\theta - f_y\sin\theta\Big) \\ &= -r\Big(\big(f_{xx}(-r\sin\theta) + f_{xy}(r\cos\theta)\big)\sin\theta + f_x\cos\theta\Big) \\ &\qquad + r\Big(\big(f_{yx}(-r\sin\theta) + f_{yy}(r\cos\theta)\big)\cos\theta - f_y\sin\theta\Big) \\ &= f_{xx}r^2\sin^2\theta - 2f_{xy}r^2\sin\theta\cos\theta + f_{yy}r^2\cos^2\theta - rf_x\cos\theta - rf_y\sin\theta \end{aligned}$$

Therefore, the punchline is

$$\begin{aligned}
\frac{1}{r}f_r + f_{rr} + \frac{1}{r^2}f_{\theta\theta} &= \frac{1}{r}\Big(f_x\cos\theta + f_y\sin\theta\Big) + \Big(f_{xx}\cos^2\theta + 2f_{xy}\sin\theta\cos\theta + f_{yy}\sin^2\theta\Big) \\
&\quad + \frac{1}{r^2}\Big(f_{xx}r^2\sin^2\theta - 2f_{xy}r^2\sin\theta\cos\theta + f_{yy}r^2\cos^2\theta - rf_x\cos\theta - rf_y\sin\theta\Big) \\
&= \cancel{\frac{1}{r}f_x\cos\theta + \frac{1}{r}}f_y\sin\theta + f_{xx}\cos^2\theta + \cancel{2f_{xy}\sin\theta\cos\theta} + f_{yy}\sin^2\theta \\
&\quad + f_{xx}\sin^2\theta - \cancel{2f_{xy}\sin\theta\cos\theta} + f_{yy}\cos^2\theta - \cancel{\frac{1}{r}f_x\cos\theta - \frac{1}{r}f_y\sin\theta} \\
&= f_{xx}\big(\sin^2\theta + \cos^2\theta\big) + f_{yy}\big(\sin^2\theta + \cos^2\theta\big) \\
&= f_{xx} + f_{yy}
\end{aligned}$$

This completes the proof. ■

Example 2.5 *Show that the* **Laplace operator** *on* $\mathbb{R}^3$,

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

has the following expression **in spherical coordinates***:*

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2}\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial \phi^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial \theta}$$

or, in even more compact notation,

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

You can check out the Plane Math page *for a derivation of this, but I would recommend that only as a way to check your work.* ■